

REINALDO FRANCISCO

**APLICAÇÃO DA TEORIA DA RESPOSTA AO ITEM (T.R.I.) NO EXAME NACIONAL
DE CURSOS (E.N.C.) DA UNICENTRO.**

**Dissertação apresentada como requisito parcial à
obtenção do grau de Mestre em Ciências, Curso
de Pós-Graduação em Métodos Numéricos em
Engenharia – Área de Concentração em
Programação Matemática, dos setores de
Tecnologia e Ciências Exatas da Universidade
Federal do Paraná.**

Orientador: Prof. Dr. Anselmo Chaves Neto

**CURITIBA
2005**

AGRADECIMENTOS

Ao professor e orientador Anselmo Chaves Neto, pelo acompanhamento e revisão do estudo, pela amizade, disponibilidade e conhecimento revelados neste período.

Aos professores: Andrea Sell Dyminski, Celso Carnieri, Jair Mendes Marques, Liliana Madalena Gramani Cumin, Maria Teresinha Arns Steiner, Neida Maria Patias Volpi, Volmir Eugênio Wilhem, pelos ensinamentos recebidos durante o curso.

A todos os colegas, pela amizade que fizemos nesta caminhada.

A todos aqueles que, direta ou indiretamente, colaboraram para a concretização deste trabalho.

SUMÁRIO

Lista de Gráficos	v
Lista de Tabelas	vi
Lista de Siglas	vii
Resumo	viii
Abstract	ix

1	INTRODUÇÃO	1
1.1	TEMA DO ESTUDO	1
1.2	OBJETIVOS	2
1.3	JUSTIFICATIVA	3
1.4	ESTRUTURA DO TRABALHO	6
2.	REVISÃO DE LITERATURA.....	7
2.1	INTRODUÇÃO À TEORIA DA RESPOSTA AO ITEM NA AVALIAÇÃO	7
2.1.1	<i>Considerações Gerais Sobre a Teoria da Resposta ao Item.</i>	9
2.2	MODELOS MATEMÁTICOS DA TEORIA DA RESPOSTA AO ITEM.....	12
2.2.1	<i>Formulação do Modelo de Rasch</i>	12
2.2.1.1	Propriedades Específicas Utilizadas no Modelo de Rasch.....	15
2.2.2	<i>Modelo Logístico de Um Parâmetro (ML1).</i>	16
2.2.3	<i>Modelo Logístico de dois Parâmetros (ML2)</i>	19
2.2.4	<i>Modelo Logístico de Três Parâmetros (ML3)</i>	21
2.2.5	<i>Interpretação dos parâmetros aplicados nas Curvas Características dos Itens</i>	23
2.3	FUNÇÃO DE INFORMAÇÃO DO TESTE $I_i(\theta)$	27
2.4	UNIDIMENSIONALIDADE E INDEPENDÊNCIA LOCAL	29
2.5	AJUSTE DO MODELO PARA UMA ÚNICA POPULAÇÃO.	30
2.5.1	<i>Estimação dos Parâmetros (a_i, b_i e c_i) pelo Método da Máxima Verossimilhança</i>	32
2.5.1.1	Estimação dos Parâmetros Considerando o Agrupamento das Proficiências (θ).....	39
2.5.1.2	Aplicação do algoritmo Newton-Raphson para a Estimação dos Parâmetros	42
2.5.1.3	Aplicação do método “Scoring” de Fisher para a Estimação dos Parâmetros.....	47
2.5.1.4	Erro-padrão para a Estimação dos Parâmetros.....	48
2.5.1.5	Escore nulo ou perfeito para a Estimação dos Parâmetros.....	48
2.5.2	<i>Estimação das habilidades/proficiências</i>	49
2.5.2.1	Aplicação do algoritmo Newton-Raphson para a estimação das proficiências	51
2.5.2.2	Aplicação do método “Scoring” de Fisher para a Estimação das Proficiências	52
2.5.2.3	Erro-padrão para a Estimação das Proficiências	53
2.5.2.4	Escore nulo ou perfeito para a Estimação das Proficiências	53
2.5.3	<i>Estimação Conjunta: Parâmetros dos itens e Proficiências</i>	54
2.5.4	<i>Máxima Verossimilhança Marginal (MVM)</i>	56
2.5.4.1	Abordagem de Bock & Lieberman	57
2.5.4.2	Métodos iterativos	62
2.5.4.3	Métodos de Quadratura Gaussiana	66
2.5.4.3.1	Equações de Estimação Conjunta em Forma de Quadratura.....	67
2.5.4.4	Abordagem de Bock & Aitkin.....	69
2.5.4.4.1	Equações de estimação em Forma de Quadratura.....	71
2.5.4.5	Aplicação do algoritmo EM.	71
2.5.5	<i>Estimação Bayesiana</i>	74
2.5.5.1	Estimação dos Parâmetros dos Itens utilizando a Metodologia Bayesiana	75

2.5.5.1.1	Distribuição a priori para a_i	77
2.5.5.1.2	Distribuição a priori para b_i	78
2.5.5.1.3	Distribuição a priori para c_i	78
2.5.5.1.4	Equações de estimação em Forma de Quadratura.....	80
2.5.5.2	Estimação Bayesiana das habilidades/proficiências.....	80
2.5.5.2.1	Estimação das Proficiências pela Moda da Posteriori – MAP.....	81
2.5.5.2.2	Estimação da Proficiência pela Média da Posteriori – EAP.....	82
3	MATERIAL E MÉTODO.....	83
3.1	MÉTODO.....	83
3.1.1	<i>Introdução ao Software BILOG.....</i>	83
3.1.1.1	Descrição das fases das análises dos dados.....	84
3.1.2	<i>Métodos para a Calibração dos Itens.....</i>	85
3.1.3	<i>Métodos Implementados para a Estimação das Proficiências.....</i>	86
3.2	MATERIAIS.....	87
4	RESULTADOS E DISCUSSÕES.....	88
4.1	RESULTADOS DOS ITENS APLICADOS NOS ENC/PROVÕES.....	88
4.1.1	<i>Análise do ENC/Provão de Matemática dos alunos da UNICENTRO - 2000.....</i>	88
4.1.1.1	<i>Análise da Confiabilidade ENCProvão 2000.....</i>	91
4.1.2	<i>Análise do ENC/Provão de Matemática dos alunos da UNICENTRO - 2001.....</i>	97
4.1.2.1	<i>Análise da Confiabilidade ENCProvão 2001.....</i>	99
4.1.3	<i>Análise do ENC/Provão de Matemática dos alunos da UNICENTRO - 2002.....</i>	106
4.1.3.1	<i>Análise da Confiabilidade ENCProvão 2002.....</i>	109
4.1.4	<i>Análise do ENC/Provão de Matemática dos alunos da UNICENTRO - 2003.....</i>	115
4.1.4.1	<i>Análise da Confiabilidade ENCProvão 2003.....</i>	117
5	CONCLUSÕES.....	125
	REFERÊNCIAS BIBLIOGRÁFICAS.....	128
	ANEXOS.....	131

LISTA DE GRÁFICOS

• Figura 1: Curvas Características de 4 Itens Típicos do ML1	18
• Figura 2: Curvas Características de 4 Itens Típicos do ML2	20
• Figura 3: Curvas Características de 6 Itens Típicos do ML3	22
• Figura 4: Modelo de Curva Característica com visualização dos Itens	23
• Figura 5: Histograma das proficiências/habilidades, ENC/Provão – 2000	89
• Figura 6: Função da Informação do Teste e Erro de Medida, Provão-2000	92
• Figura 7: Curva Característica do item 12, O MAIS DIFÍCIL	94
• Figura 8: Curva Característica do item 01, O MAIS FÁCIL	95
• Figura 9: Curva Característica do item 08, O MAIS DISCRIMINANTE	96
• Figura 10: Histograma das proficiências/habilidades, ENC/Provão – 2001	98
• Figura 11: Função da Informação do Teste e Erro de Medida, Provão -2001	101
• Figura 12: Curva Característica do item 21, O MAIS DIFÍCIL	104
• Figura 13: Curva Característica do item 08, O MAIS FÁCIL	105
• Figura 14: Curva Característica do item 03, O MAIS DISCRIMINANTE	105
• Figura 15: Histograma das proficiências/habilidades, ENC/Provão – 2002	107
• Figura 16: Função da Informação do Teste e Erro de Medida, Provão – 2002	110
• Figura 17: Curva Característica do item 20, O MAIS DIFÍCIL	112
• Figura 18: Curva Característica do item 30, O MAIS FÁCIL	113
• Figura 19: Curva Característica do item 11, O MAIS DISCRIMINANTE	114
• Figura 20: Histograma das proficiências/habilidades, ENC/Provão – 2003	116
• Figura 21: Função da Informação do Teste e Erro de Medida, Provão -2003	119
• Figura 22: Curva Característica do item 09, O MAIS DIFÍCIL	122
• Figura 23: Curva Característica do item 01, O MAIS FÁCIL	123
• Figura 24: Curva Característica do item 07, O MAIS DISCRIMINANTE	124

LISTA DE TABELAS

• Tabela 01: Relação da Quantidade de Itens x Respondentes	87
• Tabela 02: Descrição Numérica das Proficiências/Habilidades, ENC/Provão 2000.....	89
• Tabela 03: Descrição dos Acertos/Itens e Índices de Discriminação, ENC 2000.....	90
• Tabela 04: Descrição dos Escores Brutos, ENC/Provão 2000.....	91
• Tabela 05: Estimativas dos Parâmetros do Modelo TRI – ENC/Provão 2000.....	93
• Tabela 06: Média e Desvio Padrão dos Indicadores – ENC/Provão 2000.....	94
• Tabela 07: Descrição Numérica das Proficiências/Habilidades, ENC/Provão 2001.....	97
• Tabela 08: Descrição dos Acertos/Itens e Índices de Discriminação, ENC 2001.....	98
• Tabela 09: Descrição dos Escores Brutos, ENC/Provão 2001.....	100
• Tabela 10: Estimativas dos Parâmetros do Modelo TRI – ENC/Provão 2001.....	102
• Tabela 11: Média e Desvio Padrão dos Indicadores – ENC/Provão 2001.....	103
• Tabela 12: Descrição Numérica das Proficiências/Habilidades, ENC/Provão 2002.....	107
• Tabela 13: Descrição dos Acertos/Itens e Índices de Discriminação, ENC 2002.....	108
• Tabela 14: Descrição dos Escores Brutos, ENC/Provão 2002.....	109
• Tabela 15: Estimativas dos Parâmetros do Modelo TRI – ENC/Provão 2002.....	111
• Tabela 16: Média e Desvio Padrão dos Indicadores –ENC/Provão 2002.....	112
• Tabela 17: Descrição Numérica das Proficiências/Habilidades, ENC/Provão 2003.....	115
• Tabela 18: Descrição dos Acertos/Itens e Índices de Discriminação, ENC 2003.....	116
• Tabela 19: Descrição dos Escores Brutos, ENC/Provão 2003.....	118
• Tabela 20: Estimativas dos Parâmetros do Modelo TRI –ENC/Provão 2003.....	120
• Tabela 21: Média e Desvio Padrão dos Indicadores –ENC/Provão 2003.....	121
• Tabela 22: Percentual de Questões em cada Nível de Dificuldade – 2000 a 2003.....	125
• Tabela 23: Habilidade/Proficiência Média, Total de Examinados e Conceitos MEC.....	126

LISTA DE SIGLAS

CCI	Curva Característica do Item;
EAP	Esperança a Posteriori;
EMV	Estimadores de Máxima Verossimilhança;
ENADE	Exame Nacional do Desempenho dos Estudantes;
ENC	Exame Nacional de Cursos;
ENEM	Exame Nacional do Ensino Médio;
EP	Erro Padrão;
fdp	Função Densidade de Probabilidade;
IES	Instituição de Ensino Superior;
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira;
MAP	Máximo a Posteriori;
MEC	Ministério da Educação e Cultura;
ML1	Modelo Logístico de Um Parâmetro;
ML2	Modelo Logístico de Dois Parâmetros;
ML3	Modelo Logístico de Três Parâmetros;
MML	Método da Máxima Verossimilhança;
MVM	Máxima Verossimilhança Marginal;
SAEB	Sistema Nacional do Ensino Básico;
TCA	Teoria Clássica de Avaliação;
TRI	Teoria da Resposta ao Item;
UNICENTRO	Universidade Estadual do Centro-Oeste, Guarapuava-PR.

RESUMO

A avaliação quantitativa tem se tornado uma ferramenta de análise cada vez mais necessária no contexto educacional, principalmente no Brasil, onde este tipo de abordagem há pouco vem sendo implantada com técnicas mais rigorosas. Desde a Teoria Clássica de Avaliação, com métodos de aferição da qualidade do ensino através de testes, novos rumos foram dados com a Teoria da Resposta ao Item (TRI). Esta técnica tem um papel importante nessa evolução de avaliação de itens. Neste trabalho são desenvolvidos os modelos logísticos e a estimação dos parâmetros da TRI, apresentando aplicações no Exame Nacional de Cursos – ENC/Provão do curso de Matemática, realizados pelos alunos da UNICENTRO, com dados fornecidos pelo MEC/INEP. A finalidade é estudar o ganho de qualidade de resultados que o uso da TRI traz para o processo avaliatório. Como instrumento de estudo dos aspectos estatísticos dos modelos da TRI, tem também este trabalho outra importante preocupação, que é detalhar algebricamente os modelos logísticos de 1, 2 e 3 parâmetros. Para trabalhar com os dados foi utilizado o software BILOG. Os resultados obtidos apresentam considerável ganho na qualidade da avaliação com a introdução da análise de itens, através da TRI.

Palavras-chave: Avaliação, Teoria de Resposta ao Item (TRI), Teoria Clássica de Avaliação.

ABSTRACT

The quantitative evaluation has become a more and more necessary analysis tool in the educational context, mainly in Brasil, where this kind of approach has been implemented with more rigorous techniques. Since the Classical Theory of Evaluation with checking methods of learning quality through tests, new routes were given with the Item Response Theory (IRT). This technique plays an important role in this evolution of items evaluation. In this work logistical models and the estimation of parameters of some IRT models are developed, presenting applications in the Exame Nacional de Cursos – ENC/Provão of the Mathematics course, fulfilled by the students of UNICENTRO, with data supplied by MEC/INEP. The goal is to study the gain in quality that the use of IRT brings to the evaluation process. As an instrument of study of statistical aspects of IRT models, this work has also another important concern, which is to detail algebraically the logistical model of 1, 2 and 3 parameters. The software BILOG was used to work with the data. The results obtained show considerable gain in quality of evaluation, since it introduces the item analysis, through the ITR.

Key words: evaluation, Item Response Theory (IRT), Classical Theory of Evaluation

1 INTRODUÇÃO

1.1 Tema do Estudo

A partir do acompanhamento dos resultados do Exame Nacional de Cursos – ENC/Provão e da Sinopse Estatística do MEC/INEP, referente aos últimos cinco anos, observa-se uma significativa expansão do número de instituições destinadas à oferta de ensino superior no Brasil. Isto está ocorrendo desde as regras criadas pelo Ministério da Educação e Cultura (MEC) no governo Fernando Henrique Cardoso, que facilitou a criação de cursos do ensino superior.

E, com a criação do Exame Nacional do Desempenho dos Estudantes – ENADE (Lei nº 10.861, de 14/04/04), que será realizado a cada três anos, além de avaliar os estudantes, ocupar-se-á, também, com vários outros fatores ligados à instituição, tais como: condições físicas, equipamentos, corpo docente, biblioteca, currículo e ensino.

O que despertou o governo para a implantação desse sistema de avaliação de ensino superior, na verdade, foi a necessidade de formar profissionais competentes e capacitados para o desenvolvimento de novas tecnologias e para uma atuação no mercado global já que sem um sistema educacional adequado e eficaz isso torna-se mais difícil.

Sabe-se que o objetivo da avaliação do processo educacional não é apenas aquele de obter escores numéricos. Os escores correspondem apenas a uma componente na avaliação quantitativa da performance do educando. A mensuração por escores conduz à observação de outros componentes do processo de avaliação, tais como: uma análise mais aprofundada nas respostas obtidas dos avaliados às situações-problema apresentadas no Exame Nacional de Cursos - ENC/Provão, e também a verificação do nível de dificuldade encontrada pelos respondentes a partir das respostas feitas nos gabaritos.

Buscando uma análise mais detalhada das respostas, e sabendo que desde a Teoria Clássica de Avaliação os métodos de verificação da qualidade do ensino através de testes tomaram novos rumos, a técnica chamada de Teoria da Resposta ao Item (TRI) adquiriu um papel muito importante nessa evolução, sendo uma ferramenta poderosa para a avaliação educacional por possibilitar que se estabeleça a probabilidade da resposta correta dada a um item a partir da relação entre as características daquele item e da proficiência/habilidade do respondente.

Para entender como funciona esta técnica, e, como pensaram os idealizadores da Teoria da Resposta ao Item - TRI, procurou-se fazer cuidadosamente o desenvolvimento matemático dos modelos logísticos que compõem a mesma, e a aplicação no Exame Nacional de Cursos – ENC/Provão, o qual foi respondido pelos alunos formando do curso de matemática da Universidade Estadual do Centro-Oeste – UNICENTRO.

1.2 Objetivos

É claro que a exigência de conhecimentos múltiplos está relacionada à solução dos itens que compõem o Exame Nacional de Cursos - ENC/Provão do curso de Matemática da UNICENTRO, o qual será analisado do ano de 2000 a 2003. Por outro lado, cabe destacar que apesar do conceito obtido por esses alunos de matemática ter sido B no ano de 2003, nos anos anteriores o conceito era C. Esses conceitos são graduados pelo MEC/INEP em cinco faixas, delimitadas pelos percentis 12, 30, 70 e 88. O conceito C está na faixa de percentis de 30 até 70, o B de 70 até 88, e o A acima de 88. O objetivo de toda IES é buscar o conceito A, que é traduzido pelo MEC como excelência. Isto indica a necessidade de realizar-se uma análise nos itens dos ENC/Provões, a fim de ponderar-se o real grau de dificuldade dos itens dessas provas para os avaliados e verificar qual é o grau de discriminação dos itens da mesma. Neste contexto tem-se para este trabalho os seguintes objetivos:

- 1 Estudar a estimação dos parâmetros de alguns modelos da TRI, com os dados obtidos junto ao MEC/INEP, com a finalidade de verificar o desempenho dos alunos (respondentes) no Exame Nacional de Cursos – ENC/Provão, formandos de Matemática de 2000 até 2003, tendo seus escores medidos através da TRI, e destacando para cada ENC/Provão, os itens mais discriminantes, os itens mais difíceis e os itens mais fáceis de cada prova.
- 2 Detalhar matematicamente os modelos logísticos de 1, 2 e 3 parâmetros e o estudo das técnicas numéricas utilizadas na estimação dos parâmetros dos modelos TRI, para uma única população de respondentes.

1.3 Justificativa

Existem várias formas de se proceder a uma avaliação no ensino superior, podendo-se destacar algumas práticas comumente utilizadas, como as avaliações internas, representadas pelas provas, usadas pelos docentes como forma de verificar o desempenho acadêmico dos alunos; e as avaliações externas, representadas pelas avaliações institucionais, tais como o Exame Nacional do Desempenho dos Estudantes – ENADE (“Provão”), tendo como uma das finalidades verificar as proficiências básicas dos concluintes dos cursos de graduações, o exame nacional do ensino médio (ENEM) e o sistema nacional de avaliação do ensino básico (SAEB) ou ainda outros exames de avaliação estaduais (SARESP – São Paulo e AVA – Paraná) têm fornecido informações valiosas às políticas educacionais.

As avaliações fazem parte do processo de aprendizagem, mas como ressalta Perrenoud (1999, p.11), “antes de regular as aprendizagens, a avaliação regula o trabalho, as atividades, as relações de autoridade e a cooperação em aula ...” envolve a troca entre sujeitos, que deveria ser sempre construtiva e crescente. O problema da avaliação feita desta forma está no parâmetro de comparação, pois se podem cometer injustiças, quando são tratadas igualmente pessoas desiguais. E, este problema está acontecendo com as avaliações que comparam as IES públicas com as IES privadas, através dos mesmos parâmetros, pois elas recebem públicos por vezes diferentes, tanto em relação à formação quanto aos objetivos pretendidos. Mas, é lógico que qualquer destas IES formará um indivíduo e este terá todo o amparo da Constituição Federal para exercer a sua carreira profissional independente da IES que lhe habilitou.

É interessante ressaltar os objetivos das avaliações, também, em alguns países. De acordo com Yamashita (2000, p.66), nos EUA, as avaliações diversas visam basicamente a comparação entre instituições nos diferentes níveis de atividades, e as universidades competem entre si. As avaliações são promovidas pelas próprias Universidades, por associações profissionais e científicas, e até mesmo por instituições privadas, e seus resultados são considerados oportunidade de *marketing* para as Instituições de Ensino Superior - IES.

Na Europa, várias transformações têm ocorrido na Holanda, Bélgica, Alemanha, França e Inglaterra. Por exemplo, na Inglaterra o órgão de financiamento das Universidades (*University Funding Comission*) tem desenvolvido vários sistemas de avaliação para a distribuição de

recursos. Já em Portugal e na Holanda, ocorre uma espécie de auto-avaliação e não uma avaliação nacional, no modelo holandês é a própria universidade que descobre suas fraquezas para tentar melhorar seu desempenho. Vê-se que a avaliação fora do Brasil não é um processo centralizado como se tem com o MEC, o qual sabiamente está procurando criar uma cultura de avaliação.

Segundo o Ministério da Educação e Cultura – MEC (2003), os censos, os levantamentos sistemáticos de dados e os testes periódicos de avaliação têm o objetivo de oferecer um retrato da situação educacional do país e permitir que a sociedade debata e contribua com sugestões para as diretrizes do setor.

É importante lembrar que as avaliações criadas nos últimos anos pelo MEC, como o Exame Nacional de Cursos – ENC, atual Exame Nacional do Desempenho dos Estudantes – ENADE, conhecido até o ano de 2003 como “Provão”, foram colocadas para a sociedade com o objetivo principal de melhoria da qualidade da educação, dos cursos oferecidos aos estudantes.

Carvalho (2001:27-54) analisou as questões dos exames (as quais denominou itens) classificando-as em caráter instrumental, quando há a necessidade de conhecimentos específicos para resolução da questão, e de caráter formativo, exigência de conhecimentos específicos e gerais. Além disso, acrescenta a tendência do caráter instrumental que tem predominado na educação superior como um todo, homogeneizando os currículos, desconsiderando o contexto local e a multiplicidade de fatores que determinam o desempenho do estudante e valorizando a aprendizagem superficial em prejuízo da aprendizagem profunda, dentre outras consequências.

De forma que na fase do processo definida como avaliação há necessidade de medidas que verifiquem a aprendizagem dos temas ministrados na fase de execução. A eficiência do trabalho educativo está fortemente associada à verificação sistemática dos resultados alcançados. A observação sistemática das mudanças de desempenho dos educandos indica ao educador não só a realidade dos resultados do seu trabalho, como também um padrão de referencia para revisão dos seus procedimentos como educador. Uma vez que a educação sempre procura atingir objetivos planejados, são de fundamental importância verificações sistemáticas dos resultados e não fazê-las é equivalente a treinar basquetebol no escuro.

A técnica chamada de Teoria da Resposta ao Item (TRI) supõe que o desempenho em um teste pode ser explicado por características ou variáveis latentes subjacentes (não observáveis diretamente) do respondente. Estas variáveis são chamadas de proficiências ou habilidades, em

geral, procura-se reunir itens para os quais supõe-se que uma certa proficiência ou habilidade é dominante.

Mas, quando se trata de avaliação tem-se o seguinte problema, provas (exames) corrigida e classificada como excelente por um professor, são corrigidas e classificadas como medíocres por outro, esta variação decorre da falta de um padrão para correção das provas. Conseqüentemente, a prova discursiva é extremamente frágil à presença de erros e a confiabilidade (fidedignidade) desse tipo de prova é difícil de ser alcançada. H. M. Vianna (1973) cita que a chance de introdução de erros na correção de provas discursivas aumenta com a quantidade de examinandos.

Então o teste objetivo surgiu como uma tentativa de solucionar estes dois problemas falta de validade e de confiabilidade (fidedignidade) e uma opinião muito freqüente entre os pesquisadores da área de avaliação é que o teste objetivo e o teste discursivo ou de resposta livre devem ser usados para avaliar qualquer objetivo educacional. Porém, o teste objetivo é recomendado quando se tem: um número muito grande de examinandos, necessidade de graus altamente fidedignos, necessidade de rapidez nos resultados e finalmente precisa-se de avaliação imparcial, justa e válida quanto ao conteúdo examinado, segundo H. M. Vianna (1973).

O teste objetivo utilizado no Exame Nacional de Cursos – ENC/Provão, é um instrumento de medida composto por vários itens, sendo que cada um deles só admite uma resposta correta entre um número fixado de opções disponíveis. Desta forma a prova objetiva assegura julgamento justo, independente do examinador e, principalmente, hoje em dia em que a correção é feita por computador. As vantagens fundamentais do teste objetivo são as seguintes, entre outras:

- Confiabilidade (fidedignidade), o julgamento imparcial fica assegurado a todos os examinados, pois o teste pode ser corrigido até por computador. A correção do teste fica livre dos erros de julgamento que podem acontecer por problemas do examinador, tais como: estresse, mau humor ou bom humor, simpatias pessoais do examinador por aluno ou estilo de redação, letra, etc. Além de eliminar o chamado “efeito de halo”, que significa que um teste com boa apresentação, boa letra e redação fluente e agradável, termina recebendo um grau maior do que outro por conta dessas características.

- Rapidez de Correção assegura ao aluno recebimento quase imediato do resultado da avaliação, e isto conduz a uma motivação à aprendizagem.
- Repetição de questões poderá ser feita de ano para ano, desde que o número de questões seja grande. E, isto facilita o trabalho do avaliador.
- Validade fica assegurada nos testes objetivos devido a ser possível um grande número de questões, ao invés das poucas questões de uma prova discursiva. Portanto tem-se nos testes objetivos uma maior representatividade do conteúdo estudado.

Sendo que o Exame Nacional de Cursos – ENC utiliza para sua análise na grande maioria questões objetivas de múltipla escolha, onde são corrigidos os itens como certos (1) e errados (0), e para buscar uma análise mais aprofundada destas respostas será utilizado neste trabalho a técnica TRI, que tem por finalidade medir em cada item respondido a dificuldade e a discriminação dos mesmos.

1.4 Estrutura do Trabalho

Este trabalho está estruturado em cinco capítulos, sendo que a introdução aborda o tema de estudo, os objetivos, a justificativa e a estrutura do estudo. O segundo capítulo traz uma revisão aprofundada sobre a Teoria da Resposta ao Item, com a descrição matemática dos modelos logísticos de 1, 2 e 3 parâmetros, inclusive a estimação dos parâmetros dos itens e a estimação das habilidades para uma única população. Já o terceiro capítulo apresenta uma breve descrição do método utilizado que é a aplicação do software BILOG, pois para trabalhar com o material/dados coletados pelo MEC/INEP utilizou-se esse software. No quarto capítulo têm-se os resultados dos itens respondidos pelos formandos (examinados) do curso de Matemática da UNICENTRO no Exame Nacional de Cursos no período de 2000 a 2003, e as discussões sobre o emprego da TRI. No quinto e último capítulo será apresentada a conclusão final do trabalho.

2. REVISÃO DE LITERATURA

2.1 Introdução à Teoria da Resposta ao Item na Avaliação

É comum verificar-se que em processos avaliativos cuja finalidade é a seleção de candidatos ou verificação de aprendizagem são utilizados resultados obtidos em “provas” (instrumentos avaliativos de desempenho), expressos apenas por seus escores brutos ou padronizados. Isso significa dizer que, por exemplo, quanto maior a “nota” do respondente em uma prova, melhor sua classificação fato que, por se tratar de instrumentos avaliativos de medição de desempenho, não refletem o todo da referida “prova”, tão pouco o grau/índice de conhecimento do respondente sobre os temas investigados. Estatisticamente esse procedimento se caracteriza em análises e interpretações, sempre associadas ao grau obtido pelo examinado e não a um item (questão) em particular. Em termos específicos as análises e interpretações estão sempre associadas à prova como um todo; pressuposto característico da Teoria Clássica da Avaliação (TCA), conforme descrito por Vianna: “...contudo, uma característica dessa teoria consiste no fato da análise psicométrica do construto ter por ênfase o instrumento de medição (prova) como um todo e não o item” (Vianna (1973) e Gulliksen (1967)). Disto decorre que torna-se inviável a comparação entre respondentes que não foram submetidos às mesmas provas, ou pelo menos, ao que se denomina de formas paralelas de testes.

Assim procedendo, buscando-se medidas avaliativas em instrumentos de medição de desempenho, muitas questões em Educação, cujo teor avaliativo investiga aquisição qualitativa de conhecimentos, permanecem sem respostas.

Pleiteando um significativo avanço, em termos estatísticos, vislumbra-se a Teoria da Resposta ao Item (TRI) como uma metodologia de melhor validação de análises das respostas, haja vista a já citada questão da investigação qualitativa de conhecimentos, onde, parte-se do pressuposto que os itens constitutivos do instrumento avaliativo assumem a característica de “elementos centrais da análise”, quebrando o paradigma observado em análises da TCA, cuja centralização faz referência a “prova” como um todo (Valle, 1999, p.01).

Assumindo este pressuposto de validação aos itens como elementos centrais da análise, pode-se, por exemplo, comparar populações (grupos de respondentes) submetidos a provas

diferentes, considerando instrumentos investigativos que venham a abranger os mesmos temas, ou seja, que ocorram comunalidades em suas características estruturantes em termos do conteúdo a ser avaliado.

Segundo Andrade (2000, p.3-5) outra comparação possível com a utilização da TRI é aquela entre respondentes de mesmo grupo, em “provas” totalmente diferentes. Em termos práticos, suponha comparar os níveis de conhecimento entre alunos de duas séries distintas (1ª e 3ª séries do ensino médio, por exemplo). Na TCA essa comparação somente será possível caso seja aplicada a mesma prova para as duas turmas (dois grupos de respondentes). Na TRI, são necessárias apenas algumas questões em comum nas provas aplicadas às duas séries. É também, possível, por exemplo, avaliar o desenvolvimento de uma determinada série de um ano para outro, ou ainda, comparar o desempenho entre escolas públicas e privadas. Assim, várias questões de interesse prático na área da educação podem ser respondidas com a utilização da TRI, trazendo um pouco mais de informação.

Os processos avaliativos educacionais, cuja finalidade é obter resultados classificatórios podem vir a ser, por meio da TRI, compreendidos mais em termos qualitativos, do que, como o são tradicionalmente contextualizados, em termos quantitativos, mesmo porque, o reflexo de tais análises deve ser entendido como a estimação de parâmetros de investigação. E não pura e simplesmente como uma escala, como se pudéssemos fragmentar o respondente em partes iguais de zero a dez, por exemplo.

A TRI, constitui-se, portanto, em um instrumento poderoso de análise e interpretação, que ao propor modelos para os traços latentes (características que o respondente já traz com ele, e que não podem ser observadas diretamente), realiza observações de variáveis secundárias, relacionadas a estes.

A TRI propõe modelos de variáveis latentes para representar a relação entre a probabilidade de um respondente apresentar determinada resposta a um item e seus traços latentes ou proficiências na área do conhecimento avaliada, permitir, inclusive, a construção de escalas de proficiências calibradas, ou seja, permite analisar as iterações entre os respondentes e os itens.

“... talvez o aspecto mais importante da nova teoria é a promessa de fornecer medidas invariantes do desempenho cognitivo, que não dependem dos itens que compõem a prova ou das

pessoas investigadas na amostra...”. A calibração fornece a cada item, parâmetros que caracterizam suas qualidades técnicas, independentes da população investigada. “...Sendo invariantes, eles não dependem da amostra selecionada para fins de calibração. Sendo invariantes, podem ser aplicados a qualquer outra população, proporcionando resultados na mesma escala de proficiência” (Fletcher 1994, p.24).

A interpretação qualitativa sobre instrumentos de avaliação quantitativa tem se tornado cada vez mais necessária no contexto educacional, principalmente em termos de Brasil, onde este tipo de abordagem de validações qualitativas sobre instrumentos tradicionalmente quantitativos há pouco vem sendo implantada. Deste modo, a TRI avança sobremaneira há interpretação estatística, pois propicia novas condições de análises em direção à consolidação do conceito de “qualidade” em se tratando do processo de ensino e aprendizagem.

Pode-se entender por traço latente ou competências cognitivas as diferentes modalidades estruturais da inteligência que compreendem determinadas operações que o indivíduo utiliza para estabelecer relações com e entre os objetos físicos, conceitos, situações, fenômenos e pessoas. As habilidades instrumentais referem-se especificamente ao plano do saber fazer e decorrem, diretamente, do nível estrutural das competências já adquiridas e que se transformam em habilidades. Isto é, a "capacidade de agir eficazmente em um determinado tipo de situação, apoiando-se em conhecimentos, mas sem se limitar a eles" (Perrenoud, 1999).

2.1.1 Considerações Gerais Sobre a Teoria da Resposta ao Item.

Os instrumentos de avaliação de desempenho passam a ter, então, nos itens (questões, perguntas) a função de “elementos centrais”, e destes, como resultado agregados, a interpretação da “prova” / “teste” como um todo, assegurando uma validação qualitativa em tal perspectiva constitutivamente quantitativa.

Para tanto a TRI tem utilizado duas funções matemáticas para caracterizar os parâmetros métricos dos itens componentes de um teste: a função logística e a função distribuição da normal padronizada (Muñiz & Hambleton, 1992) também conhecida como ogiva Gaussiana. Ambas variam de 0 a 1 e nessa escala situa-se a probabilidade de um examinado acertar a um item específico. De modo que os modelos usados pela TRI procuram se adequar à essas funções. Cada item tem a sua Curva Característica de Informação (CCI) que segue um modelo baseado em uma

daquelas funções. As CCI's descrevem os resultados para um item em termos das avaliações dos parâmetros dos itens e, evidentemente, através das suas formas, as quais serão analisadas mais especificamente na descrição dos Modelos Matemáticos da Teoria da Resposta ao Item, no capítulo quatro de resultados e discussões.

As informações contidas nas CCIs a respeito dos parâmetros métricos dos itens, dependem do modelo teórico escolhido. Rasch, em 1960, propõe o modelo denominado “Modelo Logístico de um Parâmetro”, o qual é descrito detalhadamente na seção 2.2.2. Este modelo contém o pressuposto de que a probabilidade de acerto de um item é influenciada apenas pelo grau de dificuldade do item. O parâmetro grau de dificuldade costuma ser representado por “b”.

Um segundo modelo, denominado “Modelo Logístico de dois Parâmetros” o qual será descrito detalhadamente na seção 2.2.3, foi formulado por A. Birnbaum em 1968. Neste modelo, a probabilidade de acerto de um item é influenciada pelo grau de dificuldade “b” e pelo grau de discriminação “a”.

O terceiro modelo desenvolvido, foi denominado “Modelo Logístico de Três Parâmetros”, o qual será descrito detalhadamente na seção 2.2.4, foi construído a partir dos trabalhos de A. Birnbaum e assume-se que a probabilidade de acerto de um item é influenciada pela sua dificuldade, discriminação e probabilidade de acerto ao acaso. Conseqüentemente têm-se três parâmetros: “a”, “b” e “c”, sendo “c” a probabilidade de acerto ao acaso.

Os primeiros modelos estatísticos da TRI datam da década de 50, eles foram primeiramente desenvolvidos na forma da função distribuição da normal, ou seja,

$$\phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt.$$

Depois, foram descritos para uma forma matemática mais fácil de ser tratada, e que vem sendo usada até hoje. Essa forma é a da função logística:

$$P_i(\theta) = \frac{e^{D(\theta - b_i)}}{1 + e^{D(\theta - b_i)}},$$

que é computacionalmente mais conveniente, pois é uma função explícita dos parâmetros do item e da proficiência e não envolve cálculos mais complexos. É descrito com detalhes a partir na seção 2.2.2.

Lord, em 1952, desenvolveu o modelo unidimensional de dois parâmetros, baseado na distribuição normal acumulada (ogiva normal)

$$P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}},$$

a qual também é descrita com mais detalhes na seção 2.2.3.

Este modelo foi aplicado com as limitações computacionais da época e após algumas aplicações, sentiu-se a necessidade da incorporação de um parâmetro que tratasse do problema do acerto casual (no caso o “chute”). Então, no decorrer dos estudos, surgiu o modelo de três parâmetros

$$P(U_{ij} = 1 | \theta_j) = c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}},$$

que é descrito com mais detalhe na seção 2.2.4.

No ano de 1977, Bock & Zimowski apresentaram os modelos logísticos de 1, 2 e 3 parâmetros para duas ou mais populações de respondentes. A introdução desses modelos trouxe novas possibilidades para as comparações de rendimentos de duas ou mais populações submetidas a diferentes testes com itens comuns.

Uma das dificuldades da TRI é a estimação dos parâmetros envolvidos nos modelos, em particular quando necessita-se estimar tanto os parâmetros dos itens quanto as proficiências (θ). No começo, a estimação era feita através do método da máxima verossimilhança conjunta que envolve um numero muito grande de parâmetros a serem estimados simultaneamente e, conseqüentemente, grandes problemas computacionais. Em 1970, Bock & Lieberman introduziram o método da máxima verossimilhança marginal para a estimação dos parâmetros em duas etapas. Na primeira etapa estima-se os parâmetros dos itens, assumindo-se uma certa distribuição para as proficiências. Na segunda etapa, assumindo os parâmetros dos itens conhecidos, estimam-se as proficiências. Apesar do avanço que esse método trouxe para o problema, ele requeria que todos os parâmetros dos itens fossem estimados simultaneamente. Em 1981, Bock & Aitkin propuseram uma modificação no método acima, utilizando o algoritmo EM

de Dempster, Laird & Rubin (1977), de modo a permitir que os itens pudessem ter seus parâmetros estimados em separado, facilitando em muito o aspecto computacional do processo de estimação. Mais recentemente, métodos bayesianos foram propostos para, entre outras coisas, resolver o problema de estimação dos parâmetros dos itens respondidos corretamente ou incorretamente por todos os respondentes, e também o problema da estimação das proficiências dos respondentes que acertaram ou erraram todos os itens da prova (Andrade 2000, p.04-05).

Atualmente, a TRI vem tornando-se a técnica predominante no campo de testes em vários países, como já citado, e aqui no Brasil, a TRI foi usada pela primeira vez em 1995 na análise dos dados do Sistema Nacional de Ensino Básico – SAEB. A introdução da TRI permitiu que os desempenhos de alunos de 4^a e 8^a séries do Ensino Fundamental e de 3^a série do Ensino Fundamental pudessem ser comparadas e colocados em uma escala única de conhecimento.

2.2 Modelos Matemáticos da Teoria da Resposta ao Item

A TRI é um conjunto de modelos matemáticos que procuram representar a probabilidade de um respondente “j” dar uma resposta certa a um item “i” de um instrumento de avaliação em função dos parâmetros dos itens e do conhecimento (ou proficiências) θ_j do respondente. Quanto maior a proficiência (habilidade), maior a probabilidade de acerto do item. Segundo Valle, (1999, p. 03), os modelos propostos dependem fundamentalmente de três fatores:

- a) da natureza do item – dicotômicos ou não dicotômicos;
- b) do numero de populações envolvidas – apenas uma ou mais de uma;
- c) do número de traços latentes que estão sendo medidos – apenas um ou mais de um.

Neste trabalho são considerados os modelos logísticos que avaliam apenas um traço latente/proficiência, os chamados Modelos Unidimensionais. Existem também os modelos que consideram mais de uma proficiência, os chamados Modelos Multidimensionais.

2.2.1 Formulação do Modelo de Rasch

Trata-se de um modelo dicotômico pensado na sua forma mais simples. Prediz, por exemplo, a probabilidade condicional de um resultado binário (correto/incorreto, acerto/não-acerto), dada a competência do respondente e a dificuldade da questão.

A codificação adotada pelo matemático suíço Rasch, foi de '1' para resposta correta e '0' para resposta incorreta. O modelo, então, expressa a probabilidade de se observar uma resposta correta, ou seja, de se observar '1' ao invés de '0', como uma função da diferença entre a competência/proficiência (θ) da pessoa e a dificuldade (b) da questão. Assim, tem-se a diferença ($\theta - b$), que é a relação fundamental no modelo e que funciona como expoente na função logística. O Modelo Rasch é, portanto, uma expressão matemática para a relação entre a probabilidade de sucesso (P) e a diferença entre a habilidade/competência do examinado (θ) e a dificuldade de um item (b).

Algebricamente, tem-se com a função logística a probabilidade

$$P = \frac{\exp(\theta - b)}{1 + \exp(\theta - b)}$$

e, desenvolvendo, vem:

$$P \cdot (1 + \exp(\theta - b)) = \exp(\theta - b)$$

$$P + P \exp(\theta - b) = \exp(\theta - b)$$

$$P = \exp(\theta - b) - P \exp(\theta - b)$$

$$P = \exp(\theta - b) \cdot (1 - P)$$

$$\left(\frac{P}{1 - P} \right) = \exp(\theta - b)$$

$$\log\left(\frac{P}{1 - P}\right) = \log(\exp(\theta - b)) ,$$

então, pela propriedade dos logaritmos:

$$\log\left(\frac{P}{1 - P}\right) = \theta - b$$

Pode-se afirmar, agora, segundo Rasch que, quando a habilidade/competência é igual à dificuldade ($\theta = b$), o resultado da operação de subtrair a dificuldade (b) da habilidade (θ) é zero ($\theta - b = 0$), e ao atuar como expoente do número neperiano “e” faz com que o resultado seja um, $e^0 = 1$ (todo número elevado à zero é igual à unidade).

Então, sendo a habilidade igual à dificuldade ($\theta = b$), a probabilidade aplicada em:

$$P = \frac{\exp(\theta - b)}{1 + \exp(\theta - b)}$$

$$P = \frac{\exp(\theta - \theta)}{1 + \exp(\theta - \theta)}$$

$$P = \frac{1}{(1 + 1)}, \text{ ou}$$

$$P = \frac{1}{2} = 0,50 \times 100 = 50\%.$$

Logo a probabilidade de acertar um item quando $\theta = b$ é de 50% .

Caso a habilidade seja maior que a dificuldade ($\theta > b$) tem-se, pelo mesmo raciocínio, seguindo a formulação, resultado positivo, isto é, a probabilidade é positiva ($P > 0$), acima de zero. Caso a habilidade seja menor que a dificuldade ($\theta < b$) tem-se resultado negativo, isto é, probabilidade negativa ($P < 0$) abaixo de zero.

A unidade de medida usada por Rasch para 'calibrar' itens (estimar 'dificuldade') e medir a habilidade dos respondentes (estimar 'competência') passaram a ser chamada de 'logit' por causa da transformação logarítmica ('log odds') da probabilidade de uma resposta correta. Sempre variando com valores positivos e negativos em torno do zero arbitrário da escala.

Wright & Stone (1979, p. 17) afirmam que a competência de uma pessoa em 'logits' é o logaritmo natural da sua chance (*is the natural log odds*) de acertar questões do tipo escolhido para definir o ponto 'zero' da escala. Por outro lado, a dificuldade de um item em 'logits' é o

logaritmo natural da sua chance de induzir o não-acerto em pessoas de competência 'zero' (*is the natural log odds for eliciting failure from persons with 'zero' ability*), (Ziviani, 2002).

2.2.1.1 Propriedades Específicas Utilizadas no Modelo de Rasch

Uma suposição importante adotada nos modelos da Teoria da Resposta ao Item é a de que a resposta a uma questão não deve influenciar na resposta a outras (suposição da independência local), ou ainda, “... mantidas constantes as habilidades/proficiências, as respostas dos examinados a quaisquer dois itens são estatisticamente independentes” (Pasquali, 1996, p. 176). Correlações entre os itens seriam explicáveis somente pelo que se quer estimar nos respondentes, a competência/proficiência, ou aptidão, ou capacidade, ou traço latente freqüentemente denotada na literatura da TRI por meio da letra grega θ (teta) (correspondendo, em inglês, *a latent trait, ability, proficiency*), (Ziviani, 2002).

A possibilidade da rápida distinção entre respostas intuitivamente prováveis ou verossímeis e respostas improváveis ou inverossímeis facilita o entendimento da essência do Modelo Rasch. Intuitivamente, vê-se que a cada resposta de cada examinado pode-se fazer corresponder uma probabilidade de acerto. Pois é isso que o Modelo Rasch faz, atribui uma probabilidade de acerto à resposta para uma determinada questão dependendo apenas de dois parâmetros a serem estimados, a proficiência θ_j do respondente “j” e a dificuldade do item “i”, b_i .

A partir dessa ordenação de respondentes (pela suposta competência/proficiência) e de itens (pela suposta dificuldade) Rasch (1960, p. 66) desenvolveu um modelo matemático com a função logística para a construção de medidas baseadas na relação probabilística entre a competência da existência de apenas uma aptidão responsável pela realização de um conjunto de tarefas ou itens “...é suficiente admitir que haja uma aptidão dominante responsável pelo conjunto de itens” (Pasquali, 1996, p.176) e a dificuldade do respondente. Sobre este ponto, Muñiz (1997, p. 27-28) observa que a unidimensionalidade perfeita “... aparece como uma idealização matemática difícil, senão impossível, de se alcançar com dados psicológicos reais” e que a avaliação psicológica (traço latente do respondente) terá que se acostumar a conviver com uma unidimensionalidade imperfeita.

O Modelo de Rasch, como um modelo da TRI, centra-se na estimação conjunta da dificuldade (b_i) dos “n” itens “i”, $i = 1, 2, \dots, n$ e das proficiências dos “j” examinados (θ_j) em uma mesma escala.

Estes parâmetros, θ_j e b_i , não se definem pela pontuação observada ou pelo número de examinados que acertam o item como na Teoria Clássica de Avaliação, mas sim, pela avaliação do chamado traço latente (proficiência), e a dificuldade do item de uma forma mais geral de um domínio, da qual um teste particular seria um indicador.

Esta medição conjunta, na mesma escala, das proficiências dos examinados (θ) e da dificuldade do item (b), promove uma sensível vantagem da TRI sobre a TCA porque permite analisar as interações entre os respondentes e os itens, apresentando um diagnóstico referido à variável, identificando o tipo de situação em que um respondente (avaliando) teria alta ou baixa probabilidade de responder acertadamente.

Também traz outras vantagens como a independência de seus resultados em relação às condições com que foram obtidos (tipos de amostras ou itens). As diferenças iguais de desempenho entre os respondentes e de graus de dificuldade entre os itens têm o mesmo significado independente do ponto da escala em que se encontram (propriedades de intervalo) e os erros de medida, encontrados em qualquer processo de medição, são quantificados mais precisamente, permitindo observar-se para qual faixa de proficiência um teste é mais preciso. Detalhes deste comentário adiante na análise da figura 1 para o Modelo Logístico de um Parâmetro (ML1).

O Modelo de Rasch considera que a probabilidade de resposta correta ao item “i” (P_i) depende somente da diferença entre o nível de proficiência (conhecimento) do respondente ou examinado (θ_j) e a dificuldade do item (b_i), sendo a formulação conhecida como: Modelo Logístico de Um Parâmetro.

2.2.2 Modelo Logístico de Um Parâmetro (ML1).

Este modelo é definido pela expressão:

$$P_i(\theta) = \frac{e^{D(\theta - b_i)}}{1 + e^{D(\theta - b_i)}} .$$

Onde:

- $P_i(\theta)$: chamada de função resposta do item que é a probabilidade de um respondente escolhido ao acaso e com proficiência θ acertar o item;
- θ : nível de proficiência (conhecimento) do respondente;
- b_i : é o parâmetro que representa a dificuldade (ou de posição) do item i , medido na mesma escala da habilidade/proficiência.;
- e : base dos logaritmos neperianos;
- D : é um fator de escala usado para aproximar a função logística da ogiva Gaussiana com valor 1,7.

A expressão

$$P_i(\theta) = \frac{e^{D(\theta - b_i)}}{1 + e^{D(\theta - b_i)}},$$

é desenvolvida para a sua forma mais simples de apresentação, onde é dividindo o numerador e o denominador por $e^{D(\theta - b_i)}$, obtém-se:

$$P_i(\theta) = \frac{\frac{e^{D(\theta - b_i)}}{e^{D(\theta - b_i)}}}{\frac{1 + e^{D(\theta - b_i)}}{e^{D(\theta - b_i)}}} = \frac{1}{\frac{1}{e^{D(\theta - b_i)}} + 1}$$

Aplicando o método do MMC, vem:

$$P_i(\theta) = \frac{1}{1 + e^{-D(\theta - b_i)}},$$

agora, passando o denominador para o numerador, onde o expoente do numerador tornar-se-á negativo :

$$P_i(\theta) = \left(1 + e^{-D(\theta - b_i)}\right)^{-1}$$

que é a descrição do (ML1) na sua forma mais reduzida possível para os itens, $i = 1, 2, \dots, n$.

Essa expressão é representada graficamente na figura adiante. Esse gráfico é denominado Curva Característica do Item (CCI). A CCI representa a probabilidade de uma resposta correta (ordenada) para cada nível do construto medido (abscissa).

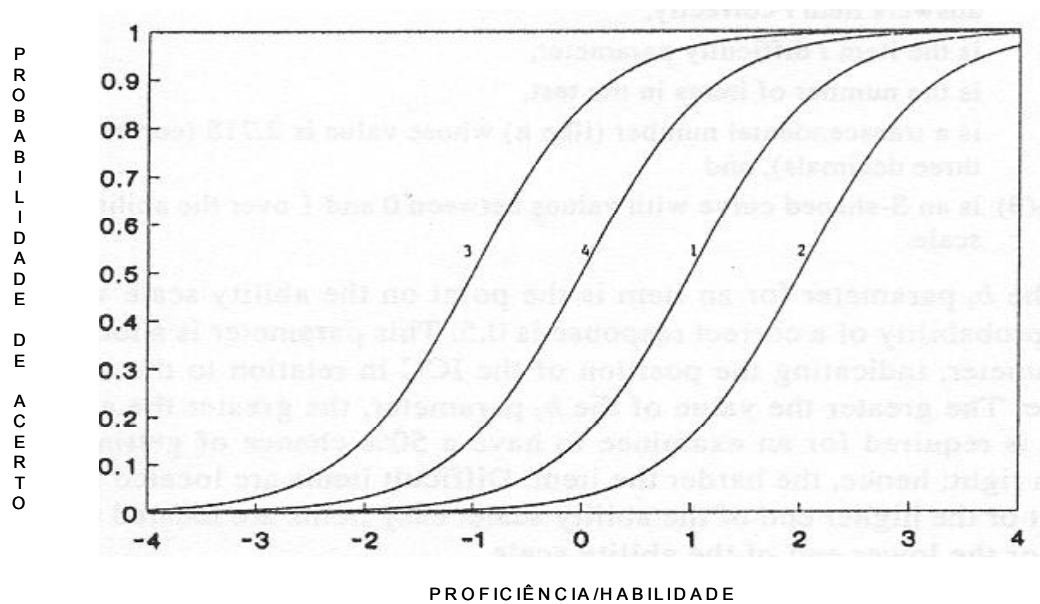


Figura 1: Curvas Características de 4 Itens Típicos do ML1 (Adaptado de Hambleton et al., 1991, p.14)

Pode-se observar na figura 1, que quando a probabilidade de resolver o item 2 (ultimo a direita) é fixada em 0,5 (valor do eixo vertical da figura), tem-se em correspondência uma proficiência/habilidade $\theta = 2$ (valor do eixo horizontal da figura). Mas, se o item considerado é o 1 (penúltimo a direita) a proficiência/habilidade necessária ao acerto diminui para $\theta = 1$, quando se considera a mesma chance de 0,5. Assim, deslocando os itens para a esquerda a proficiência/habilidade necessária diminui até $\theta = -1$. Geralmente este valor 0,5 é identificado como o grau de dificuldade limitativo “b”. De modo que no gráfico $b_3 < b_4 < b_1 < b_2$, sendo o item 3 o mais fácil e o item 2 o mais difícil.

Estes valores da abscissa podem ser expressos em diferentes métricas (Embretson; Reise, 2000). A mais utilizada é a *logit* ($\theta - b_i$) onde a diferença de uma unidade significa que o quociente entre a probabilidade de acerto e erro é igual ao número de Neper (2,7183), pois

$$\log\left(\frac{P}{1-P}\right) = \theta - b$$

$$\log\left(\frac{P}{1-P}\right) = 1$$

$$e^1 = \frac{P}{1-P} \quad .$$

Esta interpretação é a mesma em toda a escala, ou seja, ela possui propriedades de intervalo.

A localização do ponto zero (0) é arbitrária. Normalmente, no modelo de Rasch, usa-se situá-lo na média das dificuldades dos itens.

O objetivo, ao se aplicar um teste, é estimar tanto a proficiência dos examinados (θ) como a dificuldade dos itens (b). O procedimento mais usual é determinar as estimativas que tornam mais prováveis as respostas observadas.

Após a estimação de (θ) e (b), deve-se comprovar o ajuste do modelo aos dados empíricos e esta análise pode identificar itens e/ou respondentes que não se ajustam ao modelo. A ausência deste ajuste pode significar, por exemplo, itens impróprios por não serem unidimensionais ou por estarem mal formulados; também pode significar falta de conhecimento dos respondentes, etc. Cabe ao pesquisador eliminar dos resultados finais tanto aos respondentes quanto aos itens que apresentam desajuste, e, decidir por outro modelo de análise (Hambleton, 1991, pág 14).

2.2.3 Modelo Logístico de dois Parâmetros (ML2)

Este foi o primeiro modelo para TRI, Modelo Matemático Unidimensional com dois parâmetros. Foi criado por Lord em 1952 (Lord, 1967, p. 383), baseado primeiramente na função distribuição Gaussiana padronizada. A expressão desse modelo inicial é:

$$P(U_i | \theta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\theta - b_i} e^{-\frac{z^2}{2}} dz \quad ,$$

Birnbaum, mudou o suporte do modelo para a função logística. E assim a equação utilizada para avaliar a probabilidade de um examinado 'j' com proficiência θ_j responder corretamente ao i-ésimo item de um teste é dada por (Hambleton et al, 1991).

$$P(U_i | \theta_j) = \frac{e^{Da_i(\theta_j - b_i)}}{1 + e^{Da_i(\theta_j - b_i)}} \quad ,$$

utilizando-se dos mesmos passos efetuados no ML1, para simplificação, obtém-se:

$$P(U_i | \theta_j) = \frac{1}{1 + e^{-Da_i(\theta - b_i)}}$$

$$P(U_i | \theta_j) = \left[1 + e^{-Da_i(\theta - b_i)} \right]^{-1},$$

que é a sua forma mais reduzida, onde:

$P(U_i = 1|\theta)$: é a Função Característica do Item;

a_i : é o parâmetro que representa o poder de discriminação do item i , com valor proporcional à inclinação da tangente à curva no ponto com abscissa b_i ;

b_i : é o grau de dificuldade do item;

θ_j : é o nível de proficiência (conhecimento) do respondente 'j';

e : base dos logaritmos neperianos;

D : é um fator de escala usado para aproximar a função logística da ogiva Gaussiana com valor 1,7.

A figura 2 mostra as CCI do ajuste do modelo logístico para quatro itens distintos.

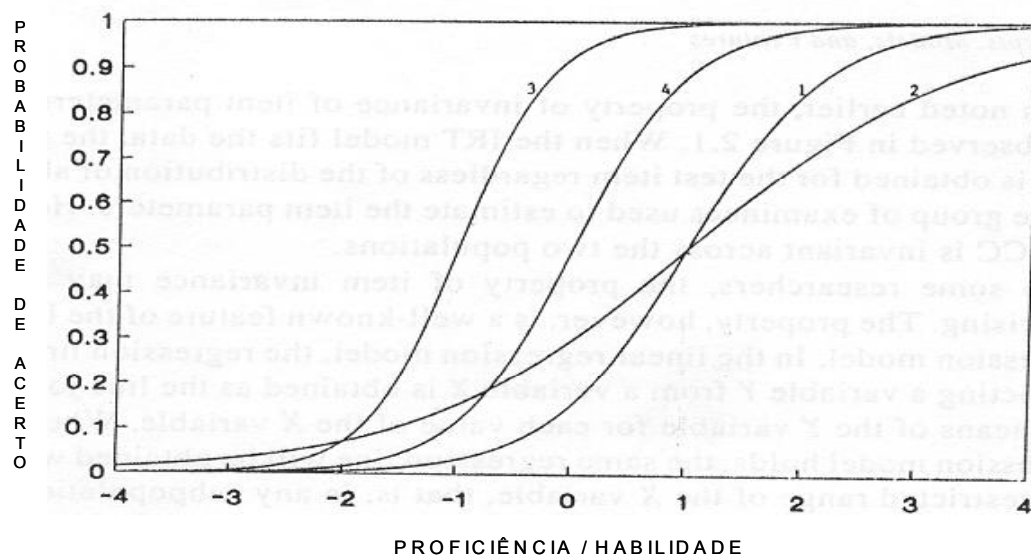


Figura 2: Curvas Características de 4 Itens Típicos do ML2 (adaptado de Hambleton et al., 1991, p.16)

O modelo logístico de dois parâmetros (ML2) é obviamente o ML1 acrescido do parâmetro índice de discriminação. Permitindo, então a discriminação dos itens. Para o item 1: $b_1 = 1$, para o item 2: $b_2 = 1$, para o item 3: $b_3 = -1$, para o item 4: $b_4 = 0$. As CCIs não são paralelas

como elas eram anteriormente no (ML1). Assim cada CCI deste modelo, tem uma inclinação diferente.

Tanto no modelo logístico de dois parâmetros como no modelo logístico de um parâmetro não é considerado que o examinando acerte o item por acaso (chute). Esta possibilidade implica em um novo parâmetro a ser incorporado no modelo.

Sabe-se que nos exames de múltipla escolha é possível o examinando responder acertadamente num item sem ter conhecimento do assunto. Assim, pensou-se no modelo com mais um parâmetro e surgiu o ML3 (Hambleton et al., 1991, p.16).

2.2.4 Modelo Logístico de Três Parâmetros (ML3)

Este modelo é denominado Modelo Logístico de Três Parâmetros, e é definido por:

$$P(U_{ij} = 1 / \theta_j) = c_i + \frac{(1 - c_i)e^{Da_i(\theta_j - b_i)}}{1 + e^{Da_i(\theta_j - b_i)}}$$

$$P(U_{ij} = 1 / \theta_j) = c_i + \frac{1 - c_i}{1 + e^{-Da_i(\theta_j - b_i)}} \quad \text{ou}$$

$$P(U_{ij} = 1 / \theta_j) = c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}}$$

$$P(U_{ij} = 1 / \theta_j) = c_i + (1 - c_i) \left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-1},$$

onde se chega a sua forma reduzida, com $i = 1, 2, 3, \dots, I$ e $j = 1, 2, 3, \dots, n$,

Onde:

$P(U_{ij} = 1/\theta_j)$: é a probabilidade do j-ésimo indivíduo escolhido ao acaso com grau de

proficiência θ_j responder corretamente ao i-ésimo item;

U_{ij} : variável dicotômica que assume o valor 1 (um) quando o j-ésimo indivíduo responde corretamente ao item i, e assume 0 (zero) quando o j-ésimo respondente não responde acertadamente ao item i;

θ_j : representa o grau de proficiência (traço latente) do j-ésimo respondente;

a_i : é o parâmetro correspondente ao índice de discriminação;

- b_i : é o parâmetro correspondente ao grau de dificuldade do item;
- c_i : é o parâmetro que representa a probabilidade de acerto ao acaso;
- e : base dos logaritmos neperianos;
- D : é um fator de escala usado para aproximar a função logística da ogiva Gaussiana com valor 1,7

A figura 3 mostra as CCI de seis ajustes do ML3 respectivamente a seis itens distintos.

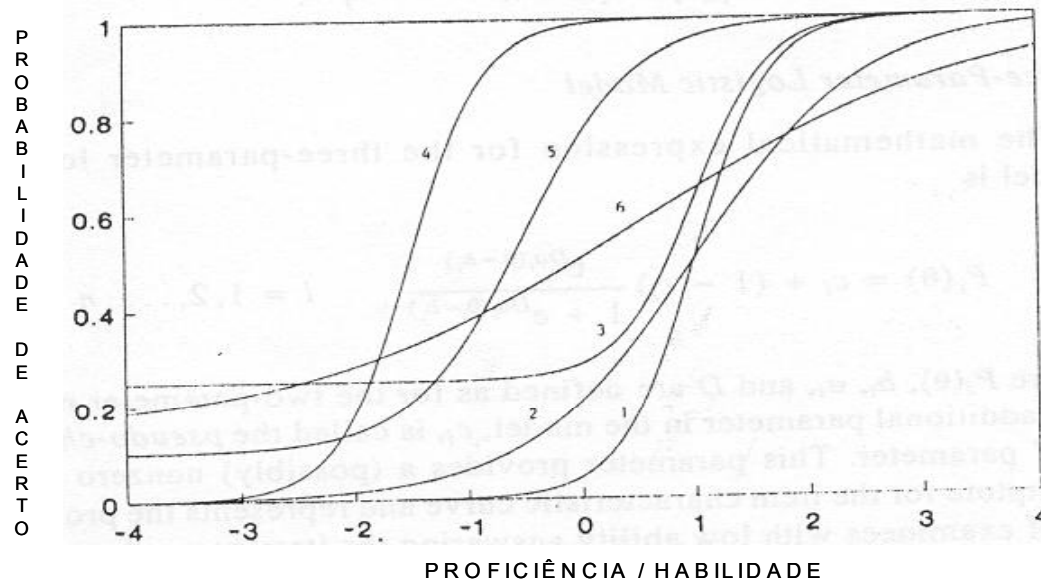


Figura 3: Curvas Características de 6 Itens Típicos do ML3 (adaptado de Hambleton et al., 1991, p.18)

Da figura 3 pode-se comparar os itens 1, 2 e 3 com 4, 5 e 6. De início, comparando o item 1 e o item 4, conclui-se que o grau de dificuldade do item 1 é muito superior ao do item 4, pois para ter uma probabilidade de 50% de acertar o item 1 necessita-se de uma proficiência perto de 1, já para o item 4 basta ter uma proficiência de perto de -2. Logo, itens mais difíceis costumam situar-se mais à direita no eixo das proficiências. Observando, agora, o item 6 nota-se que ele não está tão inclinado em relação ao eixo das abscissas, com os outros, então isto indica que ele é o menos discriminativo dos itens. Assim itens mais difíceis (Itens 1, 2 e 3) estão localizados na extremidade mais alta da escala de habilidades (a direita da origem), enquanto os itens mais fáceis estão localizados na extremidade mais baixa da escala de habilidades (a esquerda da origem). Isto, como se observou, pode ser visto claramente no gráfico da CCI.

A comparação dos itens 1 e 2 (ou itens 1, 3 e 4 “curvas mais íngremes” com itens 2, 5 e 6 “curvas mais suaves”), neste modelo o gráfico, mostra a influência do parâmetro de discriminação (a_i) na inclinação da CCI. A comparação dos itens 1 e 3 mostram a influência do parâmetro de acerto ao acaso “c” no eixo vertical desta figura, pois com uma baixíssima proficiência tem-se uma chance maior que 20% de se acertar o item.

No quarto capítulo deste trabalho, são analisados todos os itens constituintes dos Exames Nacionais de Cursos – ENC/Provão de Matemática aplicados aos alunos da UNICENTRO, no período de 2000 a 2003 e far-se-á um melhor detalhamento das CCIs.

2.2.5 Interpretação dos parâmetros aplicados nas Curvas Características dos Itens

Nota-se que $P(U_{ij} = 1|\theta_j)$ pode ser vista como a proporção de respostas corretas ao item i dentre todos os respondentes da população com habilidade/proficiência ‘ θ ’ fixada. A relação existente entre $P(U_{ij} = 1|\theta)$ e os parâmetros do modelo é mostrada na figura abaixo que é chamada de Curva Característica do Item – CCI.

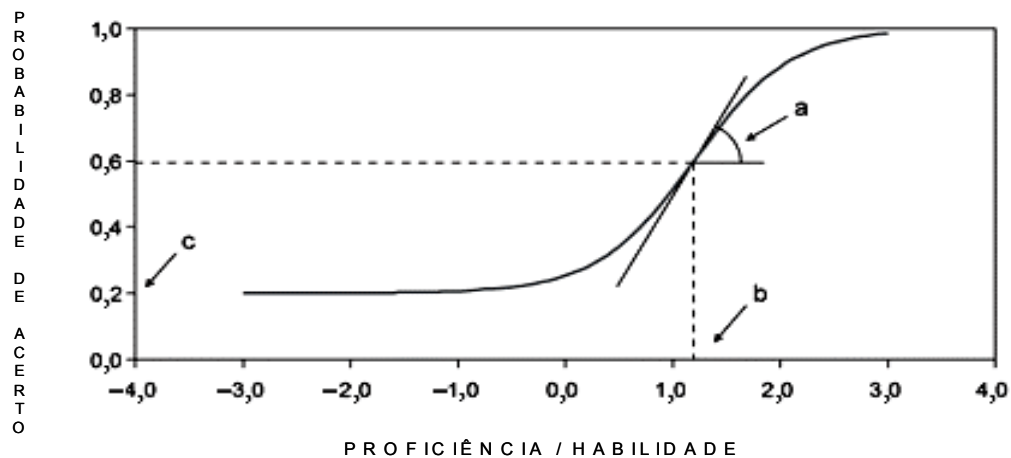


Figura 4: Modelo de Curva Característica com a visualização dos Itens.

Na área educacional o **parâmetro a** está relacionado à discriminação dos respondentes (examinados) quanto à proficiência (conhecimento) no assunto. O valor de “ a ” é proporcional à derivada da tangente à curva no ponto “ b ” (ponto de inflexão – ponto onde a curva muda de concavidade), valores negativos de “ a ” não são considerados.

Valores positivos altos do parâmetro “ a ” conduzem ao formato mais íngreme da curva, onde o item tem grande poder de discriminação, e, como consequência, a diferença entre as probabilidades das respostas corretas dos respondentes (alunos) com diferentes níveis de conhecimento (θ). Baixos valores de “ a ” indicam que o item tem pouco poder de discriminação (alunos com proficiências (θ) bastante diferentes têm aproximadamente a mesma probabilidade de responder corretamente ao item). O item não consegue discriminar indivíduos com habilidades diferentes, porém próximas. Assim à medida que o valor do parâmetro “ a ” tende a zero ($a \rightarrow 0$), a curva perde a forma de “S”, tornando-se mais achatada (mais suave). Isto significa que respondentes com diferentes níveis de habilidade/proficiência têm aproximadamente a mesma probabilidade de acertar o item, ou seja, o item não separa as diferentes proficiências.

Esse parâmetro, evidentemente, aponta a qualidade dos itens. Quanto melhor for o item, mais apropriada é a escala elaborada. Hambleton & Swaminathan (1985) afirmam que itens com $a \geq 1$ apresentam bom poder de discriminação.

Na Teoria Clássica de Avaliação - TCA, não se pensa diferente, pois, alta discriminação é entendida como uma característica desejável no item e é um indicador da qualidade deste. Uma aplicação importante que pode ser feita com o índice de discriminação é a seleção do melhor, como exemplo da discriminação de um item na TCA a qual os pesquisadores demonstram de duas maneiras:

1ª) Comparando a proporção (θ_{GS}) de examinandos que responderam ao item corretamente na parte superior do grupo (GS: escores mais altos) com a proporção (θ_{GI}) de examinandos que responderam corretamente no grupo inferior (GI: escores mais baixos). Esta comparação é feita mediante um teste estatístico de proporções e se a diferença entre as proporções for estatisticamente significativa, então o item é discriminante. O grupo superior é considerado como os 27% dos examinandos que tiveram escores mais altos no teste e o grupo inferior é composto pelos 27% dos examinados que tiveram escores mais baixos.

Nesse procedimento tem-se, então, a hipótese nula $H_0 : \theta_{GS} = \theta_{GI}$ que é testada em confronto com a hipótese alternativa $H_1 : \theta_{GS} > \theta_{GI}$. A estatística do teste de hipótese tem distribuição normal padrão, assintótica, e é dada pela expressão

$$Z = \frac{(\hat{\theta}_{GS} - \hat{\theta}_{GI}) - (\theta_{GS} - \theta_{GI})}{\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}} \sim N(0,1).$$

A diferença entre as proporções amostrais dos grupos e a diferença hipotética que figura no numerador e o erro padrão da estatística amostral no denominador. O parâmetro verdadeiro θ é estimado pela expressão

$$\hat{\theta} = \frac{n_{GS}\hat{\theta}_{GS} + n_{GI}\hat{\theta}_{GI}}{n_{GS} + n_{GI}},$$

com $n = n_{GS} + n_{GI}$, sendo n_{GS} o número de indivíduos nos grupo superior e n_{GI} o número no grupo inferior.

2ª) A maneira mais utilizada é a que calcula um coeficiente de correlação entre as variáveis: escore dos examinandos no teste (discreta: 0, 1, 2, ..., n (acertos)) e resposta dos examinandos ao item (dicotômica: 1 ou 0). O coeficiente de correlação que deve ser aplicado é o bisserial, cuja expressão é:

$$\hat{\rho}_b = \frac{\bar{X}_p - \bar{X}_t}{S_t} \cdot \frac{\hat{\theta}}{y}.$$

Onde \bar{X}_p é o escore médio entre todos os examinandos que acertaram o item i , \bar{X}_t é o escore médio global, e S_t é, o desvio padrão do teste e a proporção dos examinandos que acertaram o item i é $\hat{\theta}$; e $y = f(z)$, a ordenada na curva da Normal Padrão correspondente à área de $\hat{\theta}$.

Geralmente aplica-se este coeficiente, pois, se tem uma variável discreta e outra dicotômica.

O **parâmetro b** é medido na mesma escala do traço latente θ . Na área educacional ele está associado à dificuldade de um avaliando responder corretamente a uma questão. À medida que “b” cresce, aumenta o grau de dificuldade do item. Como já foi citado, a vantagem significativa da TRI sobre a TCA é a de que o parâmetro de dificuldade (b) e o traço latente/proficiência (θ) estão na mesma escala.

No contexto da educação o **parâmetro c** está associado à “probabilidade de um respondente (examinando) com baixa proficiência no tema do item acertá-lo”. Pode ser denominado acerto casual. O parâmetro “c” não depende da escala, é uma probabilidade e, portanto, assume valores entre 0 e 1. Tipicamente “c” assume valores que são menores do que o valor que resultaria se os examinados respondessem aleatoriamente ao item. Como Lord (1967)

notou esse fenômeno provavelmente pode ser atribuída a esperteza dos confeccionadores dos itens de uma prova, em desenvolver escolhas bastante atrativas para os respondentes, mas incorretas.

Já o **parâmetro θ** representa o traço latente/proficiência, que é uma variável que não pode ser medida diretamente. Teoricamente o parâmetro da proficiência pode assumir valores de $-\infty$ a $+\infty$. É preciso, portanto, estabelecer uma origem e uma unidade de medida para definição da escala. Estes valores são determinados de forma a representarem, respectivamente, a média (μ) e o desvio padrão (σ) das proficiências inseridas no estudo. Uma escala bastante utilizada na TRI é aquela com $\mu = 0$ e $\sigma = 1$, e, nesse caso, os valores significativos do parâmetro “b” variam (tipicamente) entre -2 e $+2$.

Exemplo seja uma escala com $\mu = 0$ e $\sigma = 1$, representada por (0,1), uma proficiência (θ) = 1,20 esta proficiência está 1,20 desvio padrão acima da proficiência média θ .

Caso fosse utilizada a escala (200, 40) e o valor da proficiência $\theta = 248$, a interpretação seria a mesma, ou seja estaria 1,20 desvio padrão acima da proficiência média.

O grande desafio está na criação de uma interpretação prática para a escala de proficiência (θ). É evidente que a interpretação fica muito fácil na escala com $\mu = 0$ e $\sigma = 1$.

Isto pode ser visto a partir da transformação de escala, voltando então a forma elaborada por Rasch e melhorada por Lord, onde:

$$a(\theta - b) = (a/40)[(40 \cdot \theta + 200) - (40 \cdot b + 200)] = a(\theta - b),$$

sendo que $a(\theta - b)$ é a parte do modelo probabilístico proposto envolvida na transformação.

Assim, segue que, para a:

$$\text{proficiência } \theta = 40 \cdot \theta + 200,$$

$$\text{dificuldade } b = 40 \cdot b + 200,$$

$$\text{discriminação } a = a/40,$$

$$\text{a probabilidade } P(U_i = 1 \mid \theta) = P(U_i = 1 \mid \theta).$$

Exemplificando, se os valores dos parâmetros “a” e “b”, na escala (0,1) são, respectivamente, $a = 0,80$ e $b = -0,20$, e seus correspondentes na escala (200; 40), e, respectivamente, $a = 0,02 = 0,80 / 40$ e $b = 192 = 40 \times (-0,20) + 200$.

Além disso, dado um respondente com proficiência $\theta = 1$, medida na escala (0, 1) tem sua proficiência representada por $\theta = 40 \times 1 + 200 = 240$ na escala (200; 40) e dado $c = 0,20$, que é o acerto ao acaso para aplicar no ML3, tem-se:

$$P(U_1 = 1 \mid \theta=1) = 0,20 + (1 - 0,20) \frac{1}{1 + e^{-1,7 \times 0,80 \times (1 - (-0,20))}} = 0,8692$$

$$P(U_1 = 1 \mid \theta=240) = 0,20 + (0,80) \cdot \frac{1}{1 + e^{-1,7 \times 0,02 \times (240 - 192)}} = 0,8692,$$

ou seja, a probabilidade de um determinado avaliado responder corretamente a um item é sempre a mesma, independentemente da escala utilizada para medir a sua proficiência (θ), ou seja, (θ) é invariante à escala de medida. Segundo Andrade (2000, p.16), não faz o menor sentido analisar itens a partir dos valores dos parâmetros de discriminação “a” e, de dificuldade “b” sem conhecer a escala de proficiência “ θ ” na qual eles foram determinados.

2.3 Função de Informação do Teste $I_i(\theta)$

Uma vez aplicado um conjunto de itens (teste) e estimado o nível de proficiência (θ) de um respondente, a TRI permite calcular o erro padrão (EP) de estimação do nível de proficiência deste respondente no qual o teste foi aplicado. Isto é uma diferença fundamental da TRI com a TCA, que assume que o erro é o mesmo para todos os examinandos.

A informação fornecida pelo teste é a soma das funções de informação dos itens em (θ). Ao analisar as Curvas Características dos Itens (CCI) uma medida bastante utilizada em conjunto é a função de Informação do Item, que é dada por:

$$I_i(\theta) = \frac{[P'_i(\theta)]^2}{P_i(\theta)Q_i(\theta)},$$

Onde : $i = 1, 2, \dots, n$.

Sendo: $I_i(\theta)$ é a “informação” fornecida pelo item “i” no nível de proficiência θ ;

$P_i(\theta)$ é a função de resposta ao item;

$$P_i(\theta) = P(X_{ij} = 1 \mid \theta) \quad \text{e} \quad Q_i(\theta) = 1 - P_i(\theta);$$

$P'_i(\theta)$ é a derivada de $P_i(\theta)$, (Birnbaum, 1968, cap. 17).

No caso do modelo logístico de 3 parâmetros, a equação de informação do item é:

$$I_i(\theta) = D^2 a_i^2 \frac{Q_i(\theta)}{P_i(\theta)} \left[\frac{P_i(\theta) - c_i}{1 - c_i} \right]^2$$

ou simplificando para parâmetros dos itens com valores já conhecidos, obtém-se a seguinte função de informação do item :

$$I_i(\theta) = \frac{2,89 a_i^2 (1 - c_i)}{\left[c_i + e^{1,7 a_i (\theta - b_i)} \right] \left[1 + e^{-1,7 a_i (\theta - b_i)} \right]^2}$$

Ela fornece a contribuição do item na estimação da proficiência, ao longo de toda a sua escala.

Sendo a função de informação do teste dada por:

$$I_i(\theta) = \sum_{i=1}^I I_i(\theta)$$

e sendo ela inversamente proporcional ao erro padrão quadrático de estimação de θ , ou seja :

$$EP = \frac{1}{\sqrt{I(\theta)}}$$

$$\sqrt{I(\theta)} = \frac{1}{EP}$$

$$I(\theta) = \frac{1}{(EP)^2}$$

A variância $V(\hat{\theta})$ afirma a importância da variação entre os valores de $\hat{\theta}$ aplicado e o valor verdadeiro de θ . Quanto menor for a variância mais confiança tem-se no teste, pois sabe-se que são poucas as diferenças entre as habilidades estimadas e os verdadeiros valores.

Quanto maior for a função de informação $I(\theta)$, menor será o erro padrão de estimação (EP), e, portanto, maior será a precisão da estimação de habilidade. O erro padrão de estimativa dá a precisão com que é estimado θ . Quanto maior o erro, menor a precisão, e seu tamanho dependem de alguns fatores, tais como:

- 1º – Do número de itens aplicados, em geral, pois, ao aumentar a quantidade de itens, diminui-se o EP,

- 2⁰– Da capacidade discriminatória dos itens, ao aumentar o parâmetro de discriminação “a” diminui-se o EP,
- 3⁰– Da diferença entre b e θ , pois quanto mais próximo está b de θ , menor será o EP.

2.4 Unidimensionalidade e Independência Local

Unidimensionalidade:

Supondo que deve haver apenas uma habilidade/proficiência (θ) responsável pela elaboração de todos os itens que compõe um teste (prova). Parece claro que qualquer desempenho humano é sempre multideterminado ou multimotivado, dado que mais de um traço latente entra na execução de qualquer tarefa. Contudo, para satisfazer o postulado da unidimensionalidade, é suficiente admitir que haja uma proficiência (θ) dominante (um fator dominante) responsável pelo conjunto de itens. Este fator é o que se supõe estar sendo medido pelo teste (Andrade 2000, p.16).

Independência Local:

Existe a independência local entre os itens de um teste, a qual será demonstrada a seguir, sendo que a resposta de um examinado a um determinado item não depende das demais respostas dadas aos outros itens.

A independência local se deriva da unidimensionalidade porque, simplesmente, significa que a resposta a um item só depende de seus parâmetros (a_i , b_i , c_i) e de “ θ ”, e não esta influenciada pela ordem de apresentação dos itens, ou pelas resposta que já tenham sido dadas, etc.

Sendo U_i a resposta do examinado para um determinado item i ($i = 1, 2, \dots, n$). Então $P(U_i | \theta)$ é a probabilidade de resposta de um examinado com proficiência θ ; $P(U_i = 1 | \theta)$ denota a probabilidade de uma resposta correta, e $P(U_i = 0 | \theta)$ denota a probabilidade de uma resposta incorreta. Sob a condição de independência local tem-se que a probabilidade de um examinando com proficiência θ acertar os n itens é:

$$\text{Probabilidade } (U_1, U_2, \dots, U_n | \theta) = P(U_1 | \theta)P(U_2 | \theta) \dots P(U_n | \theta)$$

$$P(U_1, U_2, \dots, U_n | \theta) = \prod_{i=1}^n P(U_i | \theta) .$$

A propriedade de independência local estabelece que para um determinado examinado a probabilidade de um padrão de resposta em um conjunto de itens é igual ao produto de probabilidades associado às respostas dos examinados em respostas aos itens individuais. Por

exemplo, se o padrão de resposta para um dado examinado em três itens for (1, 1, 0), isto é, $U_1 = 1$, $U_2 = 1$, e $U_3 = 0$, então:

$$P(U_1 = 1, U_2 = 1, U_3 = 0 | \theta) = P(U_1 = 1 | \theta) P(U_2 = 1 | \theta) P(U_3 = 0 | \theta) = P_1 P_2 Q_3$$

onde: $P_i = P(U_i = 1 | \theta)$ e $Q_i = 1 - P_i$ (Hambleton 1991, p.10-11).

Neste trabalho será feito o estudo dos modelos somente para itens dicotômicos.

2.5 Ajuste do Modelo para uma Única População.

A estimação dos parâmetros dos itens e das habilidades/proficiências é uma das etapas mais importantes da TRI. A probabilidade de uma resposta correta num determinado item depende da proficiência (θ) do respondente e dos parâmetros (a_i , b_i , c_i) que caracterizam o item. Mas em geral tanto os parâmetros quanto às proficiências são desconhecidas, apenas as respostas dos indivíduos aos itens do teste são conhecidas.

Na TRI, o processo de estimação dos parâmetros dos itens (a_i , b_i , c_i) é conhecido como calibração, geralmente as estimativas dos parâmetros dos itens e/ou proficiências são feitas pelo Método da Máxima Verossimilhança e também utilizando-se aplicação de algum processo iterativo, como o algoritmo Newton-Raphson e, ainda, o “Scoring” de Fischer. Alguns procedimentos bayesianos também são aplicados com bastante frequência. As soluções que exigem procedimentos iterativos, envolvem cálculos bastante complexos e, conseqüentemente, programas computacionais específicos. Em qualquer um desses casos, os valores das proficiências (θ) e dos parâmetros dos itens (a_i , b_i , c_i) estão todos na mesma escala de medida.

Quando se deseja estimar tanto os parâmetros dos itens, quanto às proficiências, há duas abordagens usuais:

- i) Estimação conjunta, dos parâmetros dos itens e proficiências,
- ii) Em duas etapas:
 - 1 Primeiro a estimação dos parâmetros dos itens e;
 - 2 Posteriormente, das proficiências.

No caso da estimação conjunta, o número de parâmetros a serem estimados simultaneamente pode ser extremamente grande ($3I + n$), sendo I o i -ésimo item e o vetor de parâmetros do item i , $i = 1, 2, \dots, I$, no caso do ML3, que é o mais utilizado. Isto leva a uma enorme exigência computacional que envolve a inversão de matrizes dessa ordem. Para contornar esse problema, Birnbaum (1968) propôs um processo fixo vai e volta (*back-and-*

forth), que é iniciado com estimativas grosseiras das habilidades (em escores padronizados, por exemplo) e envolve a estimação dos parâmetros dos itens (a_i , b_i , c_i) considerando as habilidades (θ) conhecidas; após a obtenção das estimativas dos parâmetros dos itens, as estimações das proficiências são feitas considerando conhecidos os parâmetros dos itens. Esses passos são repetidos até que algum critério de parada do processo seja alcançado. A grande vantagem desse método é que ele permite, a partir da propriedade da independência local, que os itens sejam estimados individualmente. Isto exige o trabalho com matrizes 3×3 para o ML3. De forma similar, a partir da independência entre as respostas oriundas de indivíduos diferentes, as proficiências (θ) também são estimadas individualmente. E com isso a exigência computacional diminui drasticamente.

O problema da possível inconsistência dos estimadores obtidos em uma etapa levou ao desenvolvimento da estimação em duas etapas por Bock & Lieberman (1970). Este método baseia-se na existência de uma distribuição associada à proficiência dos indivíduos da população em estudo Π . Isso possibilita que a estimação dos itens seja feita pelo Método da Máxima Verossimilhança Marginal, ou seja, considerando uma determinada distribuição para a proficiência dos indivíduos de Π , cuja função densidade de probabilidade (fdp) é $g(\theta | \eta)$, onde η é o vetor de parâmetros associados à Π e integrado a função de verossimilhança com relação à θ .

Após a estimação dos parâmetros dos itens, as proficiências são estimadas individualmente por máxima verossimilhança ou pela moda ou média da distribuição condicional de θ_j dado $u_j = (u_{j1}, \dots, u_{jI})$, o vetor de respostas do indivíduo j , $j = 1, 2, \dots, n$, com $\zeta_i = (a_i, b_i, c_i)$, o vetor de parâmetros do item i , $i = 1, 2, \dots, I$, conhecidos. Embora este método tenha a vantagem de envolver, na primeira etapa, apenas a estimação dos parâmetros dos itens, a estimação é feita através de aplicação de métodos numéricos que dependem das derivadas segundas da log-verossimilhança com relação à ζ_i e ζ_k , $i, k = 1, 2, \dots, I$, que podem ser não nulas para $i \neq k$. Com isso, há a necessidade da inversão de matrizes de ordem $3I \times 3I$ para o ML3, o que ainda pode ser bastante exigente do ponto de vista computacional.

Para contornar esse problema, Bock & Aitkin (1981) fizeram uma modificação no modelo de Bock & Lieberman adicionando a suposição de independência entre os itens, de forma que as derivadas segundas citadas acima para $i \neq k$ sejam nulas. Com isso, a matriz $3I \times 3I$ (no ML3) de derivadas segundas torna-se bloco-diagonal, o que possibilita que os (parâmetros dos) itens sejam estimados individualmente. Adicionalmente, Bock & Aitkin

propõem que a obtenção das estimativas de máxima verossimilhança seja feita com a aplicação do algoritmo EM introduzido por Dempster, Laird & Rubin (1977), (Andrade 2000, p.29).

2.5.1 Estimação dos Parâmetros (a_i , b_i , c_i) pelo Método da Máxima Verossimilhança

Considera-se agora a situação em que é conhecida a habilidade/proficiência (θ = constante) do indivíduo respondente de um determinado teste (prova) e o objetivo é estimar os parâmetros dos itens que compõem esse teste. Na prática, essa situação não ocorre, mas o estudo matemático desse caso é importante para o desenvolvimento do caso mais complexo, que é estimar os parâmetros dos itens (a_i , b_i , c_i) e das proficiências (θ) simultaneamente.

Estimar os parâmetros de todos os itens de um teste simultaneamente é um problema que envolve recursos computacionais bastante dispendiosos. Assim, na maioria dos procedimentos utilizados na TRI, os parâmetros dos itens são estimados item a item. Portanto, agora se tratando das técnicas de estimação de um único item, diga-se o item i . Para obter os parâmetros dos I itens que compõe o teste o processo deverá ser repetido independentemente I vezes.

A estimação dos parâmetros de qualquer um dos modelos logísticos da Teoria da Resposta ao Item é feita aplicando o Método da Máxima Verossimilhança (MML), quando as proficiências (θ) são conhecidas utiliza-se o (ML3) para exemplificação pois é o modelo logístico mais empregado e os (ML1) e (ML2) são derivados deste. Nas expressões a seguir assume-se que :

* As respostas dadas por examinados diferentes para determinado item são independentes;

* Examinados com a mesma proficiências (θ) respondem a todos os itens do teste de forma independente.

Assim, seja a resposta ao i -ésimo item dado pelo j -ésimo examinado representada pela variável aleatória U_{ij} que é dicotômica ou Bernoulli, ou seja, se a resposta é correta U_{ij} assume o valor '1' e se a resposta é errada assume o valor '0'. De forma que se o examinando j tem proficiência θ_j para o tema do item então:

$$U_{ij} = \begin{cases} 1 = \text{resposta correta, com } P(U_{ij} = 1) = \theta_j \\ 0 = \text{resposta incorreta, com } P(U_{ij} = 0) = 1 - \theta_j \end{cases}.$$

Então, sejam $\underline{u}_{.j} = [u_{j1}, u_{j2}, \dots, u_{jn}]$ o vetor aleatório de respostas correspondentes ao examinando j para os n itens de um teste e ainda convencionou-se denotar u_{ji} como a realização da resposta ao item i feita pelo examinando j, $\underline{u}_j = [u_{.1}, u_{.2}, \dots, u_{.n}]$ como o vetor das respostas realizadas pelo examinando j e U como a matriz do conjunto de vetores das respostas de ordem m x n. Dentro deste contexto, tem-se que a função de verossimilhança do vetor de parâmetros dos itens

$$\underline{\zeta} = [\zeta_1, \zeta_2, \dots, \zeta_I] \text{ é } L(\underline{\zeta}) = P\left(\underline{U}_{..} = \underline{u}_{..} | \theta, \underline{\zeta}\right),$$

como:

$$L(\underline{\zeta}) = \prod_{j=1}^n P(\underline{u}_{.j} = \underline{u}_{.j} | \theta_j, \underline{\zeta}) \text{ e } L(\underline{\zeta}) = \prod_{j=1}^n \prod_{i=1}^I P(u_{ji} = u_{ji} | \theta_j, \zeta_i),$$

onde se tem que a distribuição de U_{ji} só depende de ζ através de ζ_i . Usando a notação:

$$P_{ji} = P(U_{ji} = 1 | \theta_j, \zeta_i) \text{ e } Q_{ji} = 1 - P_{ji},$$

tem-se :

$$P(u_{ji} = u_{ji} | \theta_j, \zeta_i) = P(U_{ji} = 1 | \theta_j, \zeta_i)^{u_{ji}} P(U_{ji} = 0 | \theta_j, \zeta_i)^{1-u_{ji}},$$

logo:

$$P(u_{ji} = u_{ji} | \theta_j, \zeta_i) = P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}}.$$

Portanto,

$$L(\underline{\zeta}) = \prod_{j=1}^n \prod_{i=1}^I P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}},$$

é a função de verossimilhança. Aplicando o logaritmo na função de verossimilhança obtém-se a função log-verossimilhança que é o logaritmo natural da função de verossimilhança, e pode ser escrita como:

$$\log L(\underline{\zeta}) = \sum_{j=1}^n \sum_{i=1}^I \{u_{ji} \log P_{ji} + (1 - u_{ji}) \log Q_{ji}\}, \text{ (Andrade 2000, p.31-32).}$$

Que é então derivada em relação a cada um dos parâmetros: a_i , b_i e c_i . Então os Estimadores de Máxima Verossimilhança (EMV) de ζ_i , $i = 1, 2, \dots, I$, são os valores que maximizam a verossimilhança, ou equivalente, são as soluções da equação:

$$\frac{\partial \log L(\underline{\zeta})}{\partial \zeta_i} = 0, \quad i = 1, 2, \dots, I.$$

Nota-se que:

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{j=1}^n \left\{ u_{ji} \frac{\partial (\log P_{ji})}{\partial \zeta_i} + (1 - u_{ji}) \frac{\partial (\log Q_{ji})}{\partial \zeta_i} \right\}$$

e aplicando a derivada do logaritmo, tem-se:

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{j=1}^n \left\{ u_{ji} \frac{1}{P_{ji}} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right) + (1 - u_{ji}) \frac{1}{Q_{ji}} \left(\frac{\partial (1 - P_{ji})}{\partial \zeta_i} \right) \right\}$$

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{j=1}^n \left\{ u_{ji} \frac{1}{P_{ji}} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right) + (1 - u_{ji}) \frac{1}{Q_{ji}} \left(- \frac{\partial P_{ji}}{\partial \zeta_i} \right) \right\}$$

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{j=1}^n \left\{ u_{ji} \frac{1}{P_{ji}} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right) - (1 - u_{ji}) \frac{1}{Q_{ji}} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right) \right\}$$

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{j=1}^n \left\{ u_{ji} \frac{1}{P_{ji}} - (1 - u_{ji}) \frac{1}{Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right)$$

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{j=1}^n \left\{ \frac{u_{ji} Q_{ji} - P_{ji} + u_{ji} P_{ji}}{P_{ji} Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right)$$

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{j=1}^n \left\{ \frac{u_{ji} (Q_{ji} + P_{ji}) - P_{ji}}{P_{ji} Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right).$$

Sendo $Q_{ji} + P_{ji} = 1$, então:

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{j=1}^n \left\{ \frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right).$$

Por conveniência, será utilizada para as futuras demonstrações a seguinte ponderação:

$$W_{ji} = \frac{P_{ji}^* Q_{ji}^*}{P_{ji} Q_{ji}}$$

$$P_{ji} Q_{ji} = \frac{P_{ji}^* Q_{ji}^*}{W_{ji}}, \text{ (Andrade 2000, p.32).}$$

onde:

$$P_{ji}^* = \frac{1}{1 + e^{-D a_i (\theta_j - b_i)}}$$

$$P_{ji}^* = \left\{ 1 + e^{-Da_i(\theta_j - b_i)} \right\}^{-1} \text{ é a probabilidade de resposta correta.}$$

e:

$$Q_{ji}^* = 1 - P_{ji}^*$$

$$Q_{ji}^* = 1 - \left\{ 1 + e^{-Da_i(\theta_j - b_i)} \right\}^{-1}$$

$$Q_{ji}^* = \frac{e^{-Da_i(\theta_j - b_i)}}{1 + e^{-Da_i(\theta_j - b_i)}}$$

$$Q_{ji}^* = e^{-Da_i(\theta_j - b_i)} \left\{ 1 + e^{-Da_i(\theta_j - b_i)} \right\}^{-1} \text{ é a probabilidade de resposta incorreta.}$$

Então, voltando à equação

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{j=1}^n \left\{ \frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right),$$

e aplicando a ponderação, tem-se:

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{j=1}^n \left\{ \frac{u_{ji} - P_{ji}}{\frac{P_{ji}^* Q_{ji}^*}{W_{ji}}} \right\} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right)$$

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{j=1}^n \left\{ (u_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right).$$

Para obter as equações de estimação dos parâmetros a_i , b_i e c_i , precisa-se das seguintes expressões, que são derivadas do Modelo Logístico de Três Parâmetros (ML3):

$$P(U_{ji} = 1 \mid \theta_j) = c_i + (1 - c_i) \left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-1},$$

aplicando a derivada do produto no Parâmetro de Discriminação (a_i) em relação ao ML3, tem-se:

$$\frac{\partial P_{ji}}{\partial a_i} = (1 - C_i)(-1) \left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-2} e^{-D(\theta_j - b_i)} (-D(\theta_j - b_i)) + \left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-1} x_0$$

$$\frac{\partial P_{ji}}{\partial a_i} = D(1 - C_i)(\theta_j - b_i) \left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-2} e^{-D(\theta_j - b_i)},$$

sendo que:

$$P_{ji}^* = \left\{ 1 + e^{-Da_i(\theta_j - b_i)} \right\}^{-1} e$$

$$Q_{ji}^* = 1 - P_{ji}^*,$$

então para o produto:

$$P_{ji}^* Q_{ji}^* = \left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-1} \left[1 - \left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-1} \right]$$

$$P_{ji}^* Q_{ji}^* = \left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-1} \frac{e^{-Da_i(\theta_j - b_i)}}{1 + e^{-Da_i(\theta_j - b_i)}}$$

$$P_{ji}^* Q_{ji}^* = \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \cdot \frac{e^{-Da_i(\theta_j - b_i)}}{1 + e^{-Da_i(\theta_j - b_i)}}$$

$$P_{ji}^* Q_{ji}^* = \frac{e^{-Da_i(\theta_j - b_i)}}{\left(1 + e^{-Da_i(\theta_j - b_i)} \right)^2}$$

logo, substituindo na expressão onde

$$\left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-2} e^{-Da_i(\theta_j - b_i)} = P_{ji}^* Q_{ji}^*,$$

o parâmetro de discriminação (a_i) fica:

$$\frac{\partial P_{ji}}{\partial a_i} = D(1 - c_i)(\theta_j - b_i)P_{ji}^*Q_{ji}^*,$$

Aplicando a derivada do produto no ML3 para encontrar o Parâmetro de Dificuldade (b_i):

$$\frac{\partial P_{ji}}{\partial b_i} = (1 - c_i)(-1) \left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-2} e^{-D(\theta_j - b_i)} (-Da_i(0 - 1)) + \left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-1} \times 0$$

$$\frac{\partial P_{ji}}{\partial b_i} = -Da_i(1 - c_i) \left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-2} e^{-D(\theta_j - b_i)},$$

de idêntica forma como foi desenvolvida para o parâmetro de discriminação (a_i),

$$\left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-2} e^{-D(\theta_j - b_i)} = P_{ji}^*Q_{ji}^*$$

e substituindo na expressão derivada obtém-se o parâmetro de dificuldade (b_i):

$$\frac{\partial P_{ji}}{\partial b_i} = -Da_i(1 - c_i)P_{ji}^*Q_{ji}^*,$$

Aplicação da derivada do produto no ML3 para encontrar o Parâmetro de Acerto ao Acaso (c_i):

$$\frac{\partial P_{ji}}{\partial c_i} = 1 + (-1) \left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-1} + (1 - c_i) \times 0$$

$$\frac{\partial P_{ji}}{\partial c_i} = 1 + (1 - c_i)(-1) \left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-2} e^{-Da_i(\theta_j - b_i)} \cdot 0 + \left(1 - e^{-Da_i(\theta_j - b_i)} \right)^{-1} \cdot (-1)$$

$$\frac{\partial P_{ji}}{\partial c_i} = 1 - \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}},$$

aplicando o método do MMC, na expressão, vem:

$$\frac{\partial P_{ji}}{\partial c_i} = e^{-Da_i(\theta_j - b_i)} \left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-1},$$

logo:

$$\frac{\partial P_{ji}}{\partial c_i} = Q_{ji}^*.$$

Agora sim, segundo (Andrade 2000 p.33) consegue-se chegar às equações de estimação dos parâmetros dos itens:

Para obtenção do parâmetro de discriminação, (a_i):

Sendo a derivada de a_i : $\frac{\partial P_{ji}}{\partial a_i} = D(1 - c_i)(\theta_j - b_i)P_{ji}^*Q_{ji}^*$, que, substituindo na equação:

$$\frac{\partial \log L(\zeta)}{\partial a_i} = \sum_{j=1}^n \left\{ (u_{ji} - P_{ji}) \left(\frac{\partial P_{ji}}{\partial a_i} \right) \frac{W_{ji}}{P_{ji}^*Q_{ji}^*} \right\},$$

obtem-se :

$$\frac{\partial \log L(\zeta)}{\partial a_i} = \sum_{j=1}^n \left\{ (u_{ji} - P_{ji}) D(1 - c_i)(\theta_j - b_i) P_{ji}^* Q_{ji}^* \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\},$$

logo:

$$\frac{\partial \log L(\zeta)}{\partial a_i} = D(1 - c_i) \sum_{j=1}^n (u_{ji} - P_{ji})(\theta_j - b_i) W_{ji}.$$

Para obtenção do parâmetro de dificuldade, (b_i):

Dado que a derivada de b_i é : $\frac{\partial P_{ji}}{\partial b_i} = -Da_i(1 - c_i)P_{ji}^*Q_{ji}^*$, que substituindo na equação:

$$\frac{\partial \log L(\zeta)}{\partial b_i} = \sum_{j=1}^n \left\{ (u_{ji} - P_{ji}) \left(\frac{\partial P_{ji}}{\partial b_i} \right) \frac{W_{ji}}{P_{ji}^*Q_{ji}^*} \right\},$$

obtem-se :

$$\frac{\partial \log L(\zeta)}{\partial b_i} = \sum_{j=1}^n \left\{ (u_{ji} - P_{ji})(-1)Da_i(1 - c_i)P_{ji}^*Q_{ji}^* \frac{W_{ji}}{P_{ji}^*Q_{ji}^*} \right\},$$

logo :

$$\frac{\partial \log L(\zeta)}{\partial b_i} = -Da_i(1 - c_i) \sum_{j=1}^n (u_{ji} - P_{ji}) W_{ji}.$$

Para obtenção do parâmetro de acerto ao acaso, (c_i):

Sendo a derivada de c_i :

$$\frac{\partial P_{ji}}{\partial c_i} = Q_{ji}^*,$$

que substituindo em:

$$\frac{\partial \log L(\zeta)}{\partial c_i} = \sum_{j=1}^n \left\{ (u_{ji} - P_{ji}) \left(\frac{\partial P_{ji}}{\partial c_i} \right) \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\},$$

obtém-se :

$$\frac{\partial \log L(\zeta)}{\partial c_i} = \sum_{j=1}^n \left\{ (u_{ji} - P_{ji}) Q_{ji}^* \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\},$$

logo:

$$\frac{\partial \log L(\zeta)}{\partial c_i} = \sum_{j=1}^n \left\{ (u_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^*} \right\}.$$

Então, as equações de estimação para os parâmetros a_i , b_i e c_i são, respectivamente:

$$a_i : D(1 - c_i) \sum_{j=1}^n (u_{ji} - P_{ji}) (\theta_j - b_i) W_{ji} = 0$$

$$b_i : -Da_i(1 - c_i) \sum_{j=1}^n (u_{ji} - P_{ji}) W_{ji} = 0$$

$$c_i : \sum_{j=1}^n \left\{ (u_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^*} \right\} = 0$$

Este sistema não tem solução direta para a estimação dos parâmetros dos itens, mas sim por métodos iterativos, o que verificar-se-á adiante (Andrade 2000, p.34).

2.5.1.1 Estimação dos Parâmetros Considerando o Agrupamento das Proficiências (θ)

Uma outra maneira de estimação dos parâmetros dos modelos dos itens é considerar as proficiências agrupadas em “q” categorias. Isso é possível para as proficiências (θ) conhecidas, logo pode-se agrupá-las definindo um conjunto de “q” intervalos cujos valores centrais (ou alguma medida central dessas proficiências) sejam denotados por: $\bar{\theta}_k$, $k=1, 2, \dots, q$. Para fins de desenvolvimento matemático, pode-se supor que todos os indivíduos pertencentes à categoria k têm proficiências $\bar{\theta}_k$, o que pode reduzir bastante a exigência computacional tornando este método mais atrativo.

De forma geral, considera-se que “q” grupos de f_{ki} , $k=1, 2, \dots, q$, indivíduos com proficiências conhecidas $\bar{\theta}_k$ são selecionados ao acaso da população Π em estudo para responder ao item i. Seja r_{ki} o número de indivíduos do grupo k que responderam corretamente ao item i.

Pela independência local entre as respostas dos diferentes indivíduos, pode-se assumir que r_{ki} , $k=1, 2, \dots, q$, tem-se a distribuição Binomial com parâmetros f_{ki} e P_{ki} , onde P_{ki}

representa o (ML3):

$$P(U_{ji} = 1 \mid \theta_j) = C_i + (1 - C_i) \left(1 + e^{-Da_i(\theta_j - b_i)} \right)^{-1},$$

com θ_j substituída por $\bar{\theta}_k$. Com isso, a verossimilhança será:

$$L(\zeta) = \prod_{k=1}^q \prod_{i=1}^I \left\{ \binom{f_{ki}}{r_{ki}} P_{ki}^{r_{ki}} Q_{ki}^{f_{ki}-r_{ki}} \right\},$$

e a log-verossimilhança,

$$\log L(\zeta) = C + \sum_{k=1}^q \sum_{i=1}^I \left\{ r_{ki} \log P_{ki} + (f_{ki} - r_{ki}) \log Q_{ki} \right\},$$

onde:

$$C = \sum_{k=1}^q \sum_{i=1}^I \binom{f_{ki}}{r_{ki}} \text{ é constante com relação a } \zeta.$$

Tomando-se a derivada de:

$$\log L(\zeta) = C + \sum_{k=1}^q \sum_{i=1}^I \left\{ r_{ki} \log P_{ki} + (f_{ki} - r_{ki}) \log Q_{ki} \right\},$$

com relação a ζ_i , tem-se:

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{k=1}^q \left\{ r_{ki} \frac{\partial (\log P_{ki})}{\partial \zeta_i} + (f_{ki} - r_{ki}) \left(\frac{\partial \log Q_{ki}}{\partial \zeta_i} \right) \right\},$$

aplicando a derivada do logaritmo como já feito anteriormente em relação aos parâmetros dos itens, onde as proficiências não eram agrupadas mas sim constantes, tem-se:

$$\begin{aligned} \frac{\partial \log L(\zeta)}{\partial \zeta_i} &= \sum_{k=1}^q \left\{ r_{ki} \frac{1}{P_{ki}} \left(\frac{\partial P_{ki}}{\partial \zeta_i} \right) + (f_{ki} - r_{ki}) \frac{1}{Q_{ki}} \left(-\frac{\partial P_{ki}}{\partial \zeta_i} \right) \right\} \\ \frac{\partial \log L(\zeta)}{\partial \zeta_i} &= \sum_{k=1}^q \left\{ \frac{r_{ki}}{P_{ki}} - \frac{(f_{ki} - r_{ki})}{Q_{ki}} \right\} \left(\frac{\partial P_{ki}}{\partial \zeta_i} \right), \end{aligned}$$

aplicando o método do MMC, obtém-se:

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{k=1}^q \left\{ \frac{r_{ki} Q_{ki} - f_{ki} P_{ki} + r_{ki} P_{ki}}{P_{ki} Q_{ki}} \right\} \left(\frac{\partial P_{ki}}{\partial \zeta_i} \right)$$

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{k=1}^q \frac{1}{P_{ki} Q_{ki}} (r_{ki} - f_{ki} P_{ki}) \left(\frac{\partial P_{ki}}{\partial \zeta_i} \right),$$

ponderando:

$$P_{ji} Q_{ji} = \frac{P_{ji}^* Q_{ji}^*}{W_{ji}},$$

que aplicando na equação:

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{k=1}^q (r_{ki} - f_{ki} P_{ki}) \frac{W_{ki}}{P_{ki}^* Q_{ki}^*} \left(\frac{\partial P_{ki}}{\partial \zeta_i} \right).$$

Para obter as equações de estimação dos parâmetros dos itens por agrupamento das proficiências as quais foram derivadas do ML3, onde o procedimento matemático já foi visto anteriormente, utiliza-se as seguintes expressões derivadas:

$$\frac{\partial P_{ji}}{\partial a_i} = D(1 - c_i)(\theta_j - b_i) P_{ji}^* Q_{ji}^*,$$

$$\frac{\partial P_{ji}}{\partial b_i} = -D a_i (1 - c_i) P_{ji}^* Q_{ji}^*,$$

$$\frac{\partial P_{ji}}{\partial c_i} = Q_{ji}^*.$$

Para obter as equações de estimação para os parâmetros a_i , b_i e c_i , quando as proficiências são agrupadas:

$$a_i : D(1 - c_i) \sum_{k=1}^q (r_{ki} - f_{ki} P_{ki}) (\theta_k - b_i) W_{ki} = 0,$$

$$b_i : -D a_i (1 - c_i) \sum_{k=1}^q (r_{ki} - f_{ki} P_{ki}) W_{ki} = 0,$$

$$c_i : \sum_{k=1}^q (r_{ki} - f_{ki} P_{ki}) \frac{W_{ki}}{P_{ki}^*} = 0.$$

Tanto as equações de estimações anteriores, bem como estas definidas agora, não têm solução direta, mas sim por métodos iterativos para a obtenção das estimativas de máxima verossimilhança dos parâmetros dos itens. De modo que se deve aplicar o Método de Newton-

Raphson na sua forma multivariada para se obter a solução e conseqüentemente as estimativas dos parâmetros dos itens. Uma outra solução é usar-se o chamado método “*Scoring*” de Fisher. A seguir, faz-se uma breve descrição do algoritmo Newton-Raphson e do método “*Scoring*” de Fisher.

2.5.1.2 Aplicação do algoritmo Newton-Raphson para a Estimação dos Parâmetros

Segundo (Andrade 2000, p.37), seja $\ell(\zeta) = \log L(\zeta)$ a log-verossimilhança, onde $\zeta = (\zeta_1, \zeta_2, \dots, \zeta_I)$, com $\zeta_i = (a_i, b_i, c_i)^t$. Se valores iniciais $\bar{\zeta}_i^{(0)} = (a_i^{(0)}, b_i^{(0)}, c_i^{(0)})^t$ podem ser encontrados para ζ_i , então uma estimativa atualizada será

$$\bar{\zeta}_i^{(1)} = \bar{\zeta}_i^{(0)} + \Delta \bar{\zeta}_i^{(0)},$$

ou seja:

$$\hat{a}_i^{(1)} = \hat{a}_i^{(0)} + \Delta \hat{a}_i^{(0)},$$

$$\hat{b}_i^{(1)} = \hat{b}_i^{(0)} + \Delta \hat{b}_i^{(0)},$$

$$\hat{c}_i^{(1)} = \hat{c}_i^{(0)} + \Delta \hat{c}_i^{(0)},$$

onde:

$$\Delta \hat{a}_i^{(0)}, \Delta \hat{b}_i^{(0)} \text{ e } \Delta \hat{c}_i^{(0)} \text{ são erros de aproximação.}$$

Usando-se a expressão em série de Taylor de $\frac{\partial \ell(\zeta)}{\partial \zeta_i}$ em torno de $\bar{\zeta}_i^{(0)}$, tem-se:

$$\frac{\partial \ell(\zeta)}{\partial a_i} = \frac{\partial \ell(\bar{\zeta}_i^{(0)})}{\partial a_i} + \Delta \hat{a}_i^{(0)} \frac{\partial^2 \ell(\bar{\zeta}_i^{(0)})}{\partial a_i^2} + \Delta \hat{b}_i^{(0)} \frac{\partial^2 \ell(\bar{\zeta}_i^{(0)})}{\partial a_i \partial b_i} + \Delta \hat{c}_i^{(0)} \frac{\partial^2 \ell(\bar{\zeta}_i^{(0)})}{\partial a_i \partial c_i} + R_{ai}(\bar{\zeta}_i^{(0)}),$$

$$\frac{\partial \ell(\zeta)}{\partial b_i} = \frac{\partial \ell(\bar{\zeta}_i^{(0)})}{\partial b_i} + \Delta \hat{b}_i^{(0)} \frac{\partial^2 \ell(\bar{\zeta}_i^{(0)})}{\partial b_i^2} + \Delta \hat{a}_i^{(0)} \frac{\partial^2 \ell(\bar{\zeta}_i^{(0)})}{\partial b_i \partial a_i} + \Delta \hat{c}_i^{(0)} \frac{\partial^2 \ell(\bar{\zeta}_i^{(0)})}{\partial b_i \partial c_i} + R_{bi}(\bar{\zeta}_i^{(0)}),$$

$$\frac{\partial \ell(\zeta)}{\partial c_i} = \frac{\partial \ell(\bar{\zeta}_i^{(0)})}{\partial c_i} + \Delta \hat{c}_i^{(0)} \frac{\partial^2 \ell(\bar{\zeta}_i^{(0)})}{\partial c_i^2} + \Delta \hat{a}_i^{(0)} \frac{\partial^2 \ell(\bar{\zeta}_i^{(0)})}{\partial c_i \partial a_i} + \Delta \hat{b}_i^{(0)} \frac{\partial^2 \ell(\bar{\zeta}_i^{(0)})}{\partial c_i \partial b_i} + R_{ci}(\bar{\zeta}_i^{(0)}),$$

onde, $\frac{\partial \ell(\bar{\zeta}_i)}{\partial \alpha_i}$ representa a função $\frac{\partial \ell(\zeta_i)}{\partial \alpha_i}$ avaliada no ponto $\zeta_i = \bar{\zeta}_i$.

Nessas expressões utiliza-se que:

$$\frac{\partial \ell(\zeta)}{\partial \zeta_i} \text{ é função apenas de } \zeta_i, \text{ não dependendo de } \zeta_l \text{ para } l \neq i.$$

Por isso, pode-se representá-la de forma simplificada por:

$$\frac{\partial \ell(\zeta_i)}{\partial \zeta_i}.$$

Fazendo:

$$\frac{\partial l(\zeta_i)}{\partial a_i} = \frac{\partial l(\zeta_i)}{\partial b_i} = \frac{\partial l(\zeta_i)}{\partial c_i} = 0,$$

e, usando a notação:

$$\begin{aligned} L_1 &= \frac{\partial l(\bar{\zeta}_i^{(0)})}{\partial a_i} & L_{11} &= \frac{\partial^2 l(\bar{\zeta}_i^{(0)})}{\partial a_i^2} & L_{12} &= \frac{\partial^2 l(\bar{\zeta}_i^{(0)})}{\partial a_i \partial b_i} & L_{13} &= \frac{\partial^2 l(\bar{\zeta}_i^{(0)})}{\partial a_i \partial c_i}, \\ L_2 &= \frac{\partial l(\bar{\zeta}_i^{(0)})}{\partial b_i} & L_{21} &= \frac{\partial^2 l(\bar{\zeta}_i^{(0)})}{\partial b_i \partial a_i} & L_{22} &= \frac{\partial^2 l(\bar{\zeta}_i^{(0)})}{\partial b_i^2} & L_{23} &= \frac{\partial^2 l(\bar{\zeta}_i^{(0)})}{\partial b_i \partial c_i}, \\ L_3 &= \frac{\partial l(\bar{\zeta}_i^{(0)})}{\partial c_i} & L_{31} &= \frac{\partial^2 l(\bar{\zeta}_i^{(0)})}{\partial c_i \partial a_i} & L_{32} &= \frac{\partial^2 l(\bar{\zeta}_i^{(0)})}{\partial c_i \partial b_i} & L_{33} &= \frac{\partial^2 l(\bar{\zeta}_i^{(0)})}{\partial c_i^2}, \end{aligned}$$

e desprezando os restos $R_{ai}(\bar{\zeta}_i^{(0)})$, $R_{bi}(\bar{\zeta}_i^{(0)})$ e $R_{ci}(\bar{\zeta}_i^{(0)})$,

obtem-se:

$$\begin{aligned} 0 &= L_1 + L_{11} \Delta \bar{a}_i^{(0)} + L_{12} \Delta \bar{b}_i^{(0)} + L_{13} \Delta \bar{c}_i^{(0)}, \\ 0 &= L_2 + L_{12} \Delta \bar{a}_i^{(0)} + L_{22} \Delta \bar{b}_i^{(0)} + L_{23} \Delta \bar{c}_i^{(0)}, \\ 0 &= L_3 + L_{13} \Delta \bar{a}_i^{(0)} + L_{23} \Delta \bar{b}_i^{(0)} + L_{33} \Delta \bar{c}_i^{(0)}. \end{aligned}$$

Colocando o resultado obtido em forma matricial, tem-se:

$$-\begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} = \begin{pmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{pmatrix} \begin{pmatrix} \Delta \bar{a}_i^{(0)} \\ \Delta \bar{b}_i^{(0)} \\ \Delta \bar{c}_i^{(0)} \end{pmatrix}.$$

Resolvendo o sistema para $\Delta \bar{\zeta}_i^{(0)}$, obtêm-se:

$$\begin{pmatrix} \Delta \bar{a}_i^{(0)} \\ \Delta \bar{b}_i^{(0)} \\ \Delta \bar{c}_i^{(0)} \end{pmatrix} = -\begin{pmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{pmatrix}^{-1} \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix},$$

e finalmente substituindo, por

$$\begin{aligned} \bar{a}_i^{(1)} &= \bar{a}_i^{(0)} + \Delta \bar{a}_i^{(0)}, \\ \bar{b}_i^{(1)} &= \bar{b}_i^{(0)} + \Delta \bar{b}_i^{(0)}, \\ \bar{c}_i^{(1)} &= \bar{c}_i^{(0)} + \Delta \bar{c}_i^{(0)}, \end{aligned}$$

onde era:

$$\begin{pmatrix} \bar{a}_i^{(1)} \\ \bar{b}_i^{(1)} \\ \bar{c}_i^{(1)} \end{pmatrix} = \begin{pmatrix} \bar{a}_i^{(0)} \\ \bar{b}_i^{(0)} \\ \bar{c}_i^{(0)} \end{pmatrix} + \begin{pmatrix} \Delta \bar{a}_i^{(0)} \\ \Delta \bar{b}_i^{(0)} \\ \Delta \bar{c}_i^{(0)} \end{pmatrix}$$

tem-se:

$$\begin{pmatrix} \bar{a}_i^{(1)} \\ \bar{b}_i^{(1)} \\ \bar{c}_i^{(1)} \end{pmatrix} = \begin{pmatrix} \bar{a}_i^{(0)} \\ \bar{b}_i^{(0)} \\ \bar{c}_i^{(0)} \end{pmatrix} - \begin{pmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{pmatrix}^{-1} \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}.$$

Após obtido $\bar{\zeta}_i^{(1)}$, este é considerado um novo ponto inicial para a obtenção de $\bar{\zeta}_i^{(2)}$, e assim por diante. Este processo é repetido até que algum critério de parada seja alcançado. Por exemplo, até que $\Delta \bar{\zeta}_i^{(t)} = \bar{\zeta}_i^{(t)} - \bar{\zeta}_i^{(t-1)}$ seja suficientemente pequeno ou que um número pré-definido, t_{\max} , de iterações seja cumprido.

As expressões L_k , $k=1, 2, 3$ derivadas primeiras são dadas pelas equações de estimação dos parâmetros dos itens, quando as proficiências são agrupadas em “q” categorias:

$$\begin{aligned}\frac{\partial \log L(\zeta)}{\partial a_i} &= D(1 - c_i) \sum_{j=1}^n (u_{ji} - P_{ji}) (\theta_j - b_i) W_{ji} . \\ \frac{\partial \log L(\zeta)}{\partial b_i} &= -Da_i(1 - c_i) \sum_{j=1}^n (u_{ji} - P_{ji}) W_{ji} . \\ \frac{\partial \log L(\zeta)}{\partial c_i} &= \sum_{j=1}^n \left\{ (u_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}} \right\},\end{aligned}$$

e as expressões L_{kl} , $k, l = 1, 2, 3$, são as derivadas segundas, obtidas de

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i \partial \zeta_i'} = \sum_{j=1}^n \left\{ \left[\frac{\partial}{\partial \zeta_i} \left(\frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) \right] \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right)' + \left(\frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) \left(\frac{\partial^2 P_{ji}}{\partial \zeta_i \partial \zeta_i'} \right) \right\},$$

onde:

$$\left(\frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right)$$

é a derivada da função log-verossimilhança, e,

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i \partial \zeta_i'} = \sum_{j=1}^n \left\{ \left[\frac{\partial v_{ji}}{\partial \zeta_i} \right] \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right)' + v_{ji} \left(\frac{\partial^2 P_{ji}}{\partial \zeta_i \partial \zeta_i'} \right) \right\},$$

onde:

$$v_{ji} = \frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}},$$

e, aplicando-se a derivada do quociente, vem:

$$\begin{aligned}\frac{\partial v_{ji}}{\partial \zeta_i} &= \frac{\partial}{\partial \zeta_i} \left(\frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) = \frac{1}{(P_{ji} Q_{ji})^2} \left\{ -P_{ji} Q_{ji} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right) - (u_{ji} - P_{ji}) \left(\frac{\partial P_{ji} Q_{ji}}{\partial \zeta_i} \right) \right\} \\ \frac{\partial v_{ji}}{\partial \zeta_i} &= \frac{-1}{(P_{ji} Q_{ji})^2} \left\{ P_{ji} Q_{ji} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right) + (u_{ji} - P_{ji}) \left[\left(\frac{\partial P_{ji}}{\partial \zeta_i} \right) - 2P_{ji} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right) \right] \right\}\end{aligned}$$

$$\frac{\partial v_{ji}}{\partial \zeta_i} = \frac{-1}{(P_{ji} Q_{ji})^2} \left\{ P_{ji} Q_{ji} + (u_{ji} - P_{ji}) (1 - 2P_{ji}) \right\} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right),$$

$$\frac{\partial v_{ji}}{\partial \zeta_i} = \frac{-1}{(P_{ji} Q_{ji})^2} (u_{ji} - P_{ji})^2 \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right) = -v_{ji}^2 \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right).$$

A última igualdade segue do fato que $u_{ji} = u_{ji}^2$.

Considerando-se $\bar{\zeta}_i^{(t)}$ a estimativa de ζ_i na iteração t , então na iteração $t + 1$ do algoritmo Newton-Raphson tem-se que :

$$\bar{\zeta}_i^{(t+1)} = \bar{\zeta}_i^{(t)} - [H(\bar{\zeta}_i^{(t)})]^{-1} h(\bar{\zeta}_i^{(t)}),$$

onde, para chegar-se às expressões de $H(\zeta_i)$ e $h(\zeta_i)$ usada em:

$$\bar{\zeta}_i^{(t+1)} = \bar{\zeta}_i^{(t)} - [H(\bar{\zeta}_i^{(t)})]^{-1} h(\bar{\zeta}_i^{(t)}).$$

nota-se que de

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i \partial \zeta_i'} = \sum_{j=1}^n \left\{ \left[\frac{\partial v_{ji}}{\partial \zeta_i} \right] \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right)' + v_{ji} \left(\frac{\partial^2 P_{ji}}{\partial \zeta_i \partial \zeta_i'} \right) \right\},$$

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i \partial \zeta_i'} = \sum_{j=1}^n \left\{ \left(\frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) \left(\frac{\partial^2 P_{ji}}{\partial \zeta_i \partial \zeta_i'} \right) - \left(\frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right)^2 \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right) \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right)' \right\}$$

Porém:

$$\frac{\partial P_{ji}^* Q_{ji}^*}{\partial \alpha_i} = (1 - 2P_{ji}^*) \frac{\partial P_{ji}^*}{\partial \alpha_i}, \quad \alpha_i \in \{a_i, b_i, c_i\},$$

$$\frac{\partial^2 P_{ji}}{\partial \alpha_i^2} = D^2(1 - c_i) (\theta_j - b_i)^2 P_{ji}^* Q_{ji}^* (1 - 2P_{ji}^*),$$

$$\frac{\partial^2 P_{ji}}{\partial a_i \partial b_i} = -D(1 - c_i) P_{ji}^* Q_{ji}^* \left\{ 1 + D a_i (\theta_j - b_i) (1 - 2P_{ji}^*) \right\}$$

$$\frac{\partial^2 P_{ji}}{\partial a_i \partial c_i} = -D(\theta_j - b_i) P_{ji}^* Q_{ji}^*,$$

$$\frac{\partial^2 P_{ji}}{\partial b_i^2} = D^2 a_i^2 (1 - c_i) P_{ji}^* Q_{ji}^* (1 - 2P_{ji}^*) ,$$

$$\frac{\partial^2 P_{ji}}{\partial b_i \partial c_i} = D a_i P_{ji}^* Q_{ji}^* ,$$

$$\frac{\partial^2 P_{ji}}{\partial c_i^2} = \frac{\partial Q_{ji}^*}{\partial c_i} = 0.$$

Com estas expressões obtém-se a derivada segunda para os parâmetros :

$$\frac{\partial^2 P_{ji}}{\partial \zeta_i \partial \zeta_i} .$$

Sejam:

$$h_{ji} = (P_{ji}^* Q_{ji}^*)^{-1} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right) = \begin{pmatrix} D(1 - c_i)(\theta_j - b_i) \\ -D a_i (1 - c_i) \\ \frac{1}{P_{ji}^*} \end{pmatrix}, e$$

$$H_{ji} = (P_{ji}^* Q_{ji}^*)^{-1} \frac{\partial^2 P_{ji}}{\partial \zeta_i \partial \zeta_i} ,$$

$$H_{ji} = \begin{pmatrix} D^2(1 - c_i)(\theta_j - b_i)^2(1 - 2P_{ji}^*) & . & . \\ -D(1 - c_i) \left\{ 1 + D a_i (\theta_j - b_i) (1 - 2P_{ji}^*) \right\} & D^2 a_i^2 (1 - c_i) (1 - 2P_{ji}^*) & . \\ -D(\theta_j - b_i) & D a_i & 0 \end{pmatrix}$$

Com isso:

$$h(\zeta_i) \equiv \frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{j=1}^n \left\{ (u_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} (P_{ji}^* Q_{ji}^*) h_{ji} = \sum_{j=1}^n (u_{ji} - P_{ji}) W_{ji} h_{ji}.$$

$$\text{Retornando à } \frac{\partial \log L(\zeta)}{\partial \zeta_i \partial \zeta_i} = \sum_{j=1}^n \left\{ \left(\frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) \left(\frac{\partial^2 P_{ji}}{\partial \zeta_i \partial \zeta_i} \right) - \left(\frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right)^2 \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right) \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right)' \right\},$$

tem-se:

$$H(\zeta_i) \equiv \frac{\partial \log L(\zeta)}{\partial \zeta_i \partial \zeta_i} = \sum_{j=1}^n \left\{ \left(\frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) (P_{ji}^* Q_{ji}^*) h_{ji} - \left(\frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right)^2 (P_{ji}^* Q_{ji}^*)^2 h_{ji} h_{ji}' \right\}$$

$$H(\zeta_i) = \sum_{j=1}^n (u_{ji} - P_{ji}) W_{ji} \left\{ H_{ji} - (u_{ji} - P_{ji}) W_{ji} h_{ji} h'_{ji} \right\}$$

A abordagem para as proficiências agrupadas em “q” categorias, utilizando-se das expressões :

$$h(\zeta_i) = \sum_{j=1}^n (u_{ji} - P_{ji}) W_{ji} h_{ji}, e$$

$$H(\zeta_i) = \sum_{j=1}^n (u_{ji} - P_{ji}) W_{ji} \left\{ H_{ji} - (u_{ji} - P_{ji}) W_{ji} h_{ji} h'_{ji} \right\}$$

vem:

$$h(\zeta_i) = \sum_{k=1}^q (r_{ki} - f_{ki} P_{ki}) W_{ki} h_{ki},$$

$$H(\zeta_i) = \sum_{k=1}^q (r_{ki} - f_{ki} P_{ki}) W_{ki} \left\{ H_{ki} - (r_{ki} - f_{ki} P_{ki}) W_{ki} h_{ki} h'_{ki} \right\}, \text{ (Andrade 2000, p.37-41).}$$

2.5.1.3 Aplicação do método “Scoring” de Fisher para a Estimação dos Parâmetros

Para aplicação do método “Scoring” de Fisher, deve-se substituir os componentes da matriz de derivadas segundas usadas no processo iterativo de Newton-Raphson pelos seus valores esperados.

Notando que a variável aleatória U_{ji} só pode assumir dois valores:

$$U_{ji} = \begin{cases} 1, & \text{com probabilidade } P_{ji} \\ 0, & \text{com probabilidade } Q_{ji} \end{cases},$$

então U_{ji} tem distribuição Bernoulli (P_{ji}).

Segue que $E(U_{ji}) = P_{ji}$ e $E(U_{ji} - P_{ji})^2 = \text{Var}(U_{ji}) = P_{ji}Q_{ji}$. Logo, de

$$h(\zeta_i) = \sum_{k=1}^q (r_{ki} - f_{ki} P_{ki}) W_{ki} h_{ki},$$

tem-se que:

$$\Delta(\zeta_i) \equiv E(H(\zeta_i))$$

$$E(H(\zeta_i)) = \sum_{j=1}^N \left\{ E(U_{ji} - P_{ji}) W_{ji} H_{ji} - E(U_{ji} - P_{ji})^2 W_{ji}^2 h_{ji} h'_{ji} \right\}$$

$$E(H(\zeta_i)) = \sum_{j=1}^N \left\{ -P_{ji} Q_{ji} W_{ji}^2 h_{ji} h'_{ji} \right\}$$

$$E(H(\zeta_i)) = - \sum_{j=1}^n \left\{ P_{ji}^* Q_{ji}^* W_{ji} h_{ji} h'_{ji} \right\}.$$

Para as habilidades/proficiências agrupadas, a expressão acima fica:

$$\Delta(\zeta_i) = - \sum_{k=1}^q \left\{ P_{ki}^* \quad Q_{ki}^* \quad W_{ki} h_{ki}' h_{ki} \right\}.$$

A expressão para estimativa de ζ_i na iteração $t+1$ será:

$$\hat{\zeta}_i^{(t+1)} = \hat{\zeta}_i^{(t)} - [\Delta(\hat{\zeta}_i^{(t)})]^{-1} h(\hat{\zeta}_i^{(t)}). \quad (\text{Andrade 2000, p.41-42}).$$

2.5.1.4 Erro-padrão para a Estimação dos Parâmetros

Os estimadores de máxima verossimilhança gozam de propriedades assintóticas conhecidas, tais como vício nulo e eficiência. Sob algumas condições de regularidade a distribuição assintótica do estimador de máxima verossimilhança, $\hat{\zeta}_i$, é normal com vetor de média ζ_i e matriz de covariâncias dada pela inversa da matriz de informação

$$I(\zeta_i) = - E \left(\frac{\partial^2 \log L(\zeta)}{\partial \zeta_i \partial \zeta_i'} \right) = -\Delta(\zeta_i),$$

onde $\Delta(\zeta_i)$ é obtida de

$$E(H(\zeta_i)) = - \sum_{j=1}^n \left\{ P_{ji}^* \quad Q_{ji}^* \quad W_{ji} h_{ji}' h_{ji} \right\}.$$

As raízes quadradas dos elementos diagonais de $[I(\zeta_i)]^{-1}$ fornecem os erros-padrão dos estimadores \hat{a}_i , \hat{b}_i e \hat{c}_i , (Andrade 2000, p.42)

2.5.1.5 Escore nulo ou perfeito para a Estimação dos Parâmetros

Alguns problemas ocorrem na estimação por máxima verossimilhança. Se o item “i” é respondido incorretamente por todos os indivíduos, isto é, $u_{ji} = 0$, $j = 1, 2, \dots, n$, então a verossimilhança $L(\zeta) = \prod_{j=1}^n \prod_{i=1}^I P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}}$, resume-se a $L(\zeta) = \prod_{j=1}^n Q_{ji}$. Considerando os valores a_i , c_i e θ_j fixos, tem-se mudanças no valor de b_i apenas transladam Q_{ji} , sem alterar seus valores máximo e mínimo. Com isso, fixando a_i , c_i , $i = 1, 2, \dots, I$, b_l , $l \neq i$ e θ_j , o valor que maximiza a verossimilhança (EMV) será $b_i = -\infty$. Por outro lado, se o item i é respondido corretamente por todos os indivíduos, isto é, $u_{ji} = 1$, então a verossimilhança

$$L(\zeta) = \prod_{j=1}^n \prod_{i=1}^I P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}}, \text{ resume-se a } L(\zeta) = \prod_{j=1}^n P_{ji}. \text{ Com o mesmo argumento anterior,}$$

tem-se o estimador de máxima verossimilhança será $b_i = +\infty$. Problemas similares a esse ocorrem com os parâmetros a_i e c_i , (Andrade, 2000, p.42-43).

2.5.2 Estimação das habilidades/proficiências

Considera-se que os parâmetros dos itens já foram calibrados em outros testes, sendo que a calibração dos itens deve ser feita com um número grande de indivíduos, a estimação das proficiências de um grupo pequeno de indivíduos é mais confiável se forem utilizados itens já calibrados.

Nesta definição trata-se da estimação da proficiência de um único respondente, exemplo, o indivíduo 'j', e que para estimar as proficiências de 'n' indivíduos, o processo pode ser repetido independentemente n vezes.

Com os mesmos critérios utilizados para a Estimação dos Parâmetros e pela independência entre as respostas de diferentes indivíduos, a independência local, pode-se escrever o logaritmo natural da função de verossimilhança, baseada nas respostas desse j-ésimo indivíduo como:

$$\text{Log } L(\zeta) = \sum_{j=1}^n \sum_{i=1}^I \{u_{ji} \log P_{ji} + (1 - u_{ji}) \log Q_{ji}\},$$

agora como função de θ e não de ζ , ou seja:

$$\text{Log } L(\theta) = \sum_{j=1}^n \sum_{i=1}^I \left\{ u_{ji} \log P_{ji} + (1 - u_{ji}) \log Q_{ji} \right\}.$$

O Estimador de Máxima Verossimilhança - EMV para a proficiência θ_j do j-ésimo respondente é o valor de θ que maximiza a função de verossimilhança, ou equivalentemente, é solucionar a equação:

$$\frac{\partial \log L(\theta)}{\partial \theta_j} = 0, \quad j = 1, 2, \dots, n.$$

Nota-se, que:

$$\frac{\partial \log L(\theta)}{\partial \theta_j} = \sum_{i=1}^I \left\{ u_{ji} \frac{\partial (\log P_{ji})}{\partial \theta_j} + (1 - u_{ji}) \frac{\partial (\log Q_{ji})}{\partial \theta_j} \right\},$$

novamente aplicando a derivada do logaritmo, tem-se:

$$\frac{\partial \log L(\theta)}{\partial \theta_j} = \sum_{i=1}^I \left\{ u_{ji} \frac{1}{P_{ji}} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right) - (1 - u_{ji}) \frac{1}{Q_{ji}} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right) \right\} = \sum_{i=1}^I \left\{ u_{ji} \frac{1}{P_{ji}} - (1 - u_{ji}) \frac{1}{Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right)$$

$$\frac{\partial \log L(\theta)}{\partial \theta_j} = \sum_{i=1}^I \left\{ \frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right),$$

e ponderando:

$$P_{ji} Q_{ji} = \frac{P_{ji}^* Q_{ji}^*}{W_{ji}},$$

aplicando na equação, tem-se:

$$\frac{\partial \log L(\theta)}{\partial \theta_j} = \sum_{i=1}^I \left\{ (u_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right).$$

Como:

$$\frac{\partial P_{ji}}{\partial \theta_j} = D a_i (1 - c_i) P_{ji}^* Q_{ji}^*,$$

obtem-se:

$$\frac{\partial \log L(\theta)}{\partial \theta_j} = \sum_{i=1}^I \left\{ (u_{ji} - P_{ji}) D a_i (1 - c_i) P_{ji}^* Q_{ji}^* \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\}$$

$$\frac{\partial \log L(\theta)}{\partial \theta_j} = D \sum_{i=1}^I a_i (1 - c_i) (u_{ji} - P_{ji}) W_{ji}.$$

Segue então que a equação de estimação $\frac{\partial \log L(\theta)}{\partial \theta_j} = 0$, para θ_j , $j = 1, 2, \dots, n$, é:

$$\theta_j : D \sum_{i=1}^I a_i (1 - c_i) (u_{ji} - P_{ji}) W_{ji} = 0.$$

Novamente, esta equação não pode ser resolvida diretamente, pois não apresenta solução explícita para θ_j e, por isso, precisa-se de algum método iterativo para obter as estimativas desejadas. Geralmente utiliza-se dos processos iterativos Newton-Raphson e “Scoring” de Fisher.

Além disso, essa equação pode não ter solução finita para certos padrões peculiares de resposta como por exemplo, em situações de acerto total ou erro total. Nesses casos, o EMV não está definido, e a solução seria buscar outros métodos de estimação.

Como todo EMV, $\hat{\theta}_{jMV}$ (o estimador de θ_j) é normalmente distribuído com média θ_j , no caso de testes com um número suficientemente grande de itens. O erro padrão de $\hat{\theta}_{jMV}$ é dado por:

$$EP(\hat{\theta}_{jMV}) = \frac{1}{\sqrt{I(\theta_j)}},$$

onde $I(\theta_j)$ é a Função de Informação do Teste, definida na seção 2.3.

Assim pode-se observar que diferentemente do erro padrão de medida utilizado na TCA, que é constante para todos os respondentes, o EP fornecido pela TRI varia de acordo com a escala de Proficiências. Ele é tipicamente menor no centro da escala, onde geralmente há mais itens, e maior nos extremos da escala, onde há poucos itens (Valle, 1999, p. 28).

2.5.2.1 Aplicação do algoritmo Newton-Raphson para a estimação das proficiências

De forma similar ao que foi feito anteriormente na definição do algoritmo de Newton-Raphson, e considerando $\hat{\theta}_j^{(t)}$ a estimativa de θ_j na iteração t , então na iteração $t + 1$ do algoritmo Newton-Raphson tem-se que:

$$\hat{\theta}_j^{(t+1)} = \hat{\theta}_j^{(t)} - [H(\hat{\theta}_j^{(t)})]^{-1} h(\hat{\theta}_j^{(t)})$$

onde, para se chegar às expressões de $H(\theta_j)$ e $h(\theta_j)$ usadas em

$$\hat{\theta}_j^{(t+1)} = \hat{\theta}_j^{(t)} - [H(\hat{\theta}_j^{(t)})]^{-1} h(\hat{\theta}_j^{(t)}),$$

nota-se que de

$$\begin{aligned} \frac{\partial \log L(\theta)}{\partial \theta_j} &= \sum_{i=1}^I \left\{ \frac{u_{ji} - P_{ji}}{P_{ji}Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right), \\ \frac{\partial^2 \log L(\theta)}{\partial \theta_j^2} &= \sum_{i=1}^I \left\{ \left[\frac{\partial}{\partial \theta_j} \left(\frac{u_{ji} - P_{ji}}{P_{ji}Q_{ji}} \right) \right] \left(\frac{\partial P_{ji}}{\partial \theta_j} \right) + \left(\frac{u_{ji} - P_{ji}}{P_{ji}Q_{ji}} \right) \left(\frac{\partial^2 P_{ji}}{\partial \theta_j^2} \right) \right\} \\ \frac{\partial^2 \log L(\theta)}{\partial \theta_j^2} &= \sum_{i=1}^I \left\{ \left(\frac{u_{ji} - P_{ji}}{P_{ji}Q_{ji}} \right) \left(\frac{\partial^2 P_{ji}}{\partial \theta_j^2} \right) - \left(\frac{u_{ji} - P_{ji}}{P_{ji}Q_{ji}} \right)^2 \left(\frac{\partial P_{ji}}{\partial \theta_j} \right)^2 \right\} \end{aligned}$$

A segunda parcela na expressão acima é obtida por $\frac{\partial P_{ji}}{\partial \theta_j} = Da_i(1 - c_i)P_{ji}^*Q_{ji}^*$. Com

relação à primeira, tem-se:

$$\frac{\partial^2 P_{ji}}{\partial \theta_j^2} = D^2 a_i^2 (1 - c_i) P_{ji}^* Q_{ji}^* (1 - 2P_{ji}^*)$$

Sejam:

$$h_{ji} = (P_{ji}^* Q_{ji}^*)^{-1} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right) = D a_i (1 - c_i),$$

$$H_{ji} = (P_{ji}^* Q_{ji}^*)^{-1} \left(\frac{\partial^2 P_{ji}}{\partial \theta_j^2} \right) = D^2 a_i^2 (1 - c_i) (1 - 2P_{ji}^*)$$

Com isso, de $\frac{\partial \log L(\theta)}{\partial \theta_j} = D \sum_{i=1}^I a_i (1 - c_i) (u_{ji} - P_{ji}) W_{ji}$

tem-se:

$$h(\theta_j) \equiv \frac{\partial \log L(\theta)}{\partial \theta_j} = \sum_{i=1}^I \left\{ (u_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} (P_{ji}^* Q_{ji}^*) h_{ji} = \sum_{i=1}^I (u_{ji} - P_{ji}) W_{ji} h_{ji}.$$

Retornando agora à

$$\frac{\partial \log L(\theta)}{\partial \theta_j} = \sum_{i=1}^I \left\{ \left(\frac{u_{ji} - P_{ji}}{P_{ji}^* Q_{ji}^*} \right) \left(\frac{\partial P_{ji}}{\partial \theta_j} \right) - \left(\frac{u_{ji} - P_{ji}}{P_{ji}^* Q_{ji}^*} \right)^2 \left(\frac{\partial P_{ji}}{\partial \theta_j} \right)^2 \right\},$$

$$H(\theta_j) \equiv \frac{\partial^2 \log L(\theta)}{\partial \theta_j^2} = \sum_{i=1}^I \left\{ \left(\frac{u_{ji} - P_{ji}}{P_{ji}^* Q_{ji}^*} \right) (P_{ji}^* Q_{ji}^*) H_{ji} - \left(\frac{u_{ji} - P_{ji}}{P_{ji}^* Q_{ji}^*} \right)^2 (P_{ji}^* Q_{ji}^*)^2 h_{ji}^2 \right\}$$

$$H(\theta_j) \equiv \sum_{i=1}^I (u_{ji} - P_{ji}) W_{ji} \left\{ H_{ji} - (u_{ji} - P_{ji}) W_{ji} h_{ji}^2 \right\}$$

com:

$$h_{ji} = (P_{ji}^* Q_{ji}^*)^{-1} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right) = D a_i (1 - c_i), \text{ e,}$$

$$H_{ji} = (P_{ji}^* Q_{ji}^*)^{-1} \left(\frac{\partial^2 P_{ji}}{\partial \theta_j^2} \right) = D^2 a_i^2 (1 - c_i) (1 - 2P_{ji}^*)$$

2.5.2.2 Aplicação do método “Scoring” de Fisher para a Estimação das Proficiências

Para aplicação do método “Scoring” de Fisher, deve-se substituir os componentes da matriz de derivadas segundas usadas no processo iterativo de Newton-Raphson pelos seus valores esperados.

Por:

$$H(\theta_j) \equiv \sum_{i=1}^I (u_{ji} - P_{ji}) W_{ji} \left\{ H_{ji} - (u_{ji} - P_{ji}) W_{ji} h_{ji}^2 \right\}$$

tem-se que:

$$\begin{aligned} \Delta(\theta_j) &\equiv E(H(\theta_j)) = \sum_{i=1}^I \left\{ E(U_{ji} - P_{ji}) W_{ji} H_{ji} - E(U_{ji} - P_{ji})^2 W_{ji}^2 h_{ji}^2 \right\} \\ \Delta(\theta_j) &\equiv E(H(\theta_j)) = - \sum_{i=1}^I P_{ji}^* Q_{ji}^* W_{ji} h_{ji}^2. \end{aligned}$$

Neste caso, a expressão para estimativa de θ_j , $j = 1, 2, \dots, n$ na iteração $t + 1$ será:

$$\hat{\theta}_j^{(t+1)} = \hat{\theta}_j^{(t)} - [\Delta(\hat{\theta}_j^{(t)})]^{-1} h(\hat{\theta}_j^{(t)}).$$

2.5.2.3 Erro-padrão para a Estimação das Proficiências

Sob algumas condições de regularidade a distribuição assintótica do estimador de máxima verossimilhança, $\hat{\theta}_j$ é normal com média θ_j e variância dada pela inversa da matriz de informação:

$$I(\theta_j) = -E \left(\frac{\partial^2 \log L(\theta)}{\partial \theta_j^2} \right) = -\Delta(\theta_j),$$

onde $\Delta(\theta_j)$ é obtida de

$$\Delta(\theta_j) = - \sum_{i=1}^I P_{ji}^* Q_{ji}^* W_{ji} h_{ji}^2.$$

A raiz quadrada de $I(\theta_j)$ fornece o erro-padrão de $\hat{\theta}_j$.

2.5.2.4 Escore nulo ou perfeito para a Estimação das Proficiências

Tal como na estimação dos parâmetros dos itens, existe um problema a ser contornado na estimação por máxima verossimilhança. Se o indivíduo j obtém escore nulo, isto é, $u_{ji} = 0$, $i = 1, 2, \dots, I$, então a verossimilhança resume-se a $L(\theta) = \prod_{i=1}^I Q_{ji}$. Como Q_{ji} , $i = 1, 2, \dots, I$, é decrescente com θ_j , então $L(\theta)$ também é decrescente com θ_j e daí o estimador de máxima verossimilhança será $\theta_j = -\infty$. Por outro lado, se o indivíduo j obter o escore total, isto é, $u_{ji} = 1$, $i = 1, 2, \dots, I$, então a verossimilhança resume-se a $L(\theta) = \prod_{i=1}^I P_{ji}$. Como P_{ji} , $i = 1, 2, \dots, I$, é crescente com θ_j , então $L(\theta)$ também é crescente com θ_j e daí o estimador de máxima verossimilhança será $\theta_j = +\infty$, (Andrade 2000, p.46-48).

2.5.3 Estimação Conjunta: Parâmetros dos itens e Proficiências

Agora tem-se o caso mais comum, em que nem os parâmetros dos itens e nem as habilidades/proficiências são conhecidos. Nas expressões já discriminadas:

$$a_i : \quad D(1 - c_i) \sum_{j=1}^n (u_{ji} - P_{ji}) (\theta_j - b_i) W_{ji} = 0$$

$$b_i : \quad -Da_i(1 - c_i) \sum_{j=1}^n (u_{ji} - P_{ji}) W_{ji} = 0$$

$$c_i : \quad \sum_{j=1}^n \left\{ (u_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^*} \right\} = 0, \text{ e,}$$

$$\theta_j : \quad D \sum_{i=1}^I a_i (1 - c_i) (u_{ji} - P_{ji}) W_{ji} = 0,$$

têm-se as equações de estimação para os parâmetros dos itens e proficiências. Entretanto, estas equações não apresentam expressões explícitas para os respectivos EMV. Por conta disso, algum processo iterativo deve ser aplicado no processo de maximização e, como consequência, algumas quantidades ou suposições podem ser adicionadas ao modelo.

A principal diferença da estimação conjunta se dá no tratamento da métrica (escala) em que todos os parâmetros são estimados. Na estimação conjunta não há uma métrica definida e, portanto, deve-se estabelecê-la. A explicação formal para a necessidade do estabelecimento da métrica dos parâmetros consiste em um problema denominado falta de identificabilidade do modelo. Essa não-identificabilidade ocorre porque mais de um conjunto de parâmetros produz o mesmo valor no ML3, e conseqüentemente, na verossimilhança.

$$\text{Se: } \theta_j^* = \alpha \theta_j + \beta,$$

$$b_i^* = \alpha b_i + \beta,$$

$$a_i^* = \frac{a_i}{\alpha} \text{ e}$$

$$c_i^* = c_i,$$

onde α e β são constantes reais com $\alpha > 0$, então o ML3:

$$\begin{aligned}
P(U_{ji} = 1 | \theta_j^*, \zeta_i^*) &= c_i^* + (1 - c_i^*) \left[1 + \exp \left[-Da_i^* (\theta_j^* - b_i^*) \right] \right]^{-1} \\
P(U_{ji} = 1 | \theta_j^*, \zeta_i^*) &= c_i + (1 - c_i) \left\{ 1 + \exp \left[-D \frac{a_i}{\alpha} (\alpha \theta_j + \beta - (\alpha b_i + \beta)) \right] \right\}^{-1} \\
P(U_{ji} = 1 | \theta_j^*, \zeta_i^*) &= c_i + (1 - c_i) \left[1 + \exp \left[-Da_i (\theta_j - b_i) \right] \right]^{-1} = P(U_{ji} = 1 | \theta_j, \zeta_i)
\end{aligned}$$

Essa não-identificabilidade pode ser eliminada de várias formas, como fixando alguns valores para as habilidades/proficiências, por exemplo. Entretanto, deve-se ressaltar que essa não-identificabilidade está intimamente relacionada à características da população envolvida no estudo. Até agora não foi especificado quando uma proficiência pode ser considerada alta ou baixa, nem como diagnosticar o quanto uma proficiência está afastada de outra. Isso pode ser resolvido especificando uma medida de posição (média, por exemplo) e outra de dispersão (desvio-padrão, por exemplo) para as proficiências. Dessa forma se está definindo uma métrica (unidade de medida) para as proficiências e, conseqüentemente, para os parâmetros dos itens. Elimina-se então o problema de não-identificabilidade do modelo padronizando as proficiências de forma que estas tenham uma média especificada μ e desvio-padrão σ . Desta forma, as proficiências e os parâmetros dos itens são estimados na métrica (μ, σ) . Em muitas situações adota-se $\mu = 0$ e $\sigma = 1$.

Para aplicação do algoritmo Newton-Raphson são necessárias as derivadas segundas da log-verossimilhança, com relação a ζ_i e θ_j , $i = 1, 2, \dots, I$ e $j = 1, 2, \dots, n$. Estas derivadas compõem uma matriz H quadrada de ordem $(3I + n)$ e essa dimensão pode ser suficientemente grande de forma a causar uma enorme exigência computacional. Por isso, precisa-se explorar um pouco mais a estrutura de H . Nota-se que pela independência local, onde:

$$\frac{\partial L(\zeta, \theta)}{\partial \zeta_i \partial \zeta_\ell} = 0, \text{ para } i \neq \ell.$$

Pela independência entre as respostas de indivíduos diferentes, tem-se:

$$\frac{\partial L(\zeta, \theta)}{\partial \theta_j \partial \theta_\ell} = 0, \text{ para } j \neq \ell.$$

Vale notar que

$$\frac{\partial L(\zeta, \theta)}{\partial \zeta_i \partial \zeta'_i} = 0 \text{ e } \frac{\partial L(\zeta, \theta)}{\partial \theta_j \partial \theta_\ell} = 0,$$

são conseqüências das suposições inerentes do modelo. Uma suposição adicional que simplifica bastante a estrutura de H é a de que não existe correlação entre itens e proficiências. Essa suposição condiz com situações práticas, pois as proficiências são inerentes dos indivíduos, que em nada dependem dos itens envolvidos no estudo. Como conseqüência desta suposição, tem-se:

$$\frac{\partial L(\zeta, \theta)}{\partial \zeta_i \partial \theta_j} = 0,$$

para $i = 1, 2, \dots, I$ e $j = 1, 2, \dots, n$.

Assim, a matriz H torna-se bloco-diagonal, na qual os “I” primeiros blocos são matrizes 3x3 relativas aos parâmetros dos itens e os n blocos seguintes são escalares relativos às proficiências. As expressões $\frac{\partial L(\zeta, \theta)}{\partial \zeta_i \partial \zeta'_i} = 0$, a $\frac{\partial L(\zeta, \theta)}{\partial \zeta_i \partial \theta_j} = 0$, facilitam bastante a estrutura de H, mas não diminuem sua dimensão. Entretanto, com base nessa estrutura bloco-diagonal, Birnbaum (1968) propôs um algoritmo em que os itens e as proficiências são estimados individualmente, utilizando o algoritmo Newton-Raphson ou o método “Scoring” de Fisher, no qual cada iteração é composta de dois estágios:

Estágio 1: Começando com estimativas iniciais para as proficiências θ (escores padronizados, por exemplo) e tratando estas proficiências como conhecidas, estima-se ζ_i , $i = 1, 2, \dots, I$.

Estágio 2: Começando com estimativas iniciais (obtidas no Estágio 1) para ζ e tratando estes parâmetros como conhecidos, estima-se as proficiências θ_j , $j = 1, 2, \dots, n$.

Este processo de dois estágios é repetido até a convergência das proficiências e dos parâmetros dos itens, (Andrade 2000, p.49-51).

2.5.4 Máxima Verossimilhança Marginal (MVM)

O método da Máxima Verossimilhança Marginal, proposto por Bock & Lieberman (1970) apresenta algumas vantagens em relação ao método da Máxima Verossimilhança Conjunta. A proposta desse método é fazer a estimação em duas etapas: primeiro os parâmetros dos itens e,

posteriormente, as proficiências. Como as proficiências não são conhecidas, precisa-se usar algum artifício de forma que a verossimilhança não seja mais função das proficiências. Anderson (1980) argumenta que considera-se uma população Π composta por n indivíduos com proficiências θ_j , $j = 1, 2, \dots, n$, e constrói-se a distribuição de frequência acumulada:

$$G(\theta) = \frac{\left(\text{número de } j: \theta_j \leq \theta \right)}{n},$$

então, se n for suficientemente grande os θ_j estarão bastante próximos de forma que $G(\theta)$ pode ser aproximada por uma distribuição contínua. A densidade $g(\theta)$, relativa à $G(\theta)$, pode realmente ser considerada a função densidade para θ no experimento de retirar um indivíduo ao acaso da população Π e observar seu parâmetro θ .

Um artifício para eliminar as proficiências na verossimilhança consiste em marginalizar a verossimilhança integrando-a com relação à distribuição da proficiência. De forma geral, considera-se que as proficiências, θ_j , $j = 1, 2, \dots, n$, são realizações de uma variável aleatória θ com distribuição contínua e função densidade de probabilidade (fdp) $g(\theta|\eta)$, duplamente diferenciável, com as componentes de η conhecidas e finitas. Para o caso em que θ tem distribuição Normal, tem-se $\eta = (\mu, \sigma^2)$, onde μ é a média e σ^2 a variância das proficiências dos indivíduos de Π . Portanto, para que os itens sejam estimados na métrica (0,1), deve-se adotar $\mu = 0$ e $\sigma = 1$. (Andrade 2000, p.52).

2.5.4.1 Abordagem de Bock & Lieberman

Com as notações acima, tem-se que a probabilidade marginal de U_j é dada por:

$$P(u_j|\zeta, \eta) = \int_{IR} P(u_j|\theta, \zeta, \eta) g(\theta|\eta) d\theta = \int_{IR} P(u_j|\theta, \zeta) g(\theta|\eta) d\theta,$$

onde na última igualdade usou-se que a distribuição de U_j não é função de η e \Re representa o conjunto dos números reais. Usando a independência entre as respostas de diferentes indivíduos, pode-se escrever a probabilidade associada ao vetor de respostas $U..$ como:

$$P(u..|\zeta, \eta) = \prod_{j=1}^n P(u_j|\zeta, \eta).$$

Embora a verossimilhança possa ser escrita como $P(u_{.j}|\zeta, \eta) = \prod_{j=1}^n P(u_{j.}|\zeta, \eta)$, tem sido freqüente utilizar a abordagem de Padrões de Resposta. Como existem I itens no total, com 2 possíveis respostas para cada item (0 ou 1), há $S = 2^I$ possíveis respostas (padrões de resposta). Quando o número de indivíduos é grande em relação ao número de itens, pode haver vantagens computacionais em trabalhar com o número de ocorrências dos diferentes padrões de resposta. O índice j não mais representará um indivíduo, mas sim um padrão de resposta.

Seja r_j o número de ocorrências distintas do padrão de resposta j , e ainda $s \leq \min(n, S)$ o número de padrões de resposta com $r_j > 0$. Segue disso que:

$$\sum_{j=1}^s r_j = n.$$

Pela independência entre as respostas dos diferentes indivíduos, têm-se os dados seguem uma distribuição Multinomial, isto é:

$$L(\zeta, \eta) = \frac{n!}{\prod_{j=1}^s r_j!} \prod_{j=1}^s [P(u_{j.}|\zeta, \eta)]^{r_j},$$

e, portanto, a log-verossimilhança é:

$$\log L(\zeta, \eta) = \log \left\{ \frac{n!}{\prod_{j=1}^s r_j!} \right\} + \sum_{j=1}^s r_j \log P(u_{j.}|\zeta, \eta)$$

As equações de estimação para os parâmetros dos itens são dadas por:

$$\frac{\partial \log L(\zeta, \eta)}{\partial \zeta_i} = 0, \quad i = 1, 2, \dots, I,$$

com:

$$\frac{\partial \log L(\zeta, \eta)}{\partial \zeta_i} = \frac{\partial}{\partial \zeta_i} \left\{ \sum_{j=1}^s r_j \log P(u_{j.}|\zeta, \eta) \right\} = \sum_{j=1}^s r_j \frac{1}{P(u_{j.}|\zeta, \eta)} \frac{\partial P(u_{j.}|\zeta, \eta)}{\partial \zeta_i}.$$

Mas:

$$\begin{aligned} \frac{\partial P(u_{j.}|\zeta, \eta)}{\partial \zeta_i} &= \frac{\partial}{\partial \zeta_i} \int_{IR} P(u_{j.}|\theta, \zeta) g(\theta|\eta) d\theta = \int_{IR} \left(\frac{\partial}{\partial \zeta_i} P(u_{j.}|\theta, \zeta) \right) g(\theta|\eta) d\theta \\ \frac{\partial P(u_{j.}|\zeta, \eta)}{\partial \zeta_i} &= \int_{IR} \left(\frac{\partial}{\partial \zeta_i} \prod_{l=1}^I P(u_{jl}|\theta, \zeta_l) \right) g(\theta|\eta) d\theta \end{aligned}$$

$$\frac{\partial P(u_{j.}|\zeta, \eta)}{\partial \zeta_i} = \int_{IR} \left(\prod_{l \neq i} P(u_{jl}|\theta, \zeta_l) \right) \left(\frac{\partial}{\partial \zeta_i} P(u_{ji}|\theta, \zeta_i) \right) g(\theta|\eta) d\theta$$

$$\frac{\partial P(u_{j.}|\zeta, \eta)}{\partial \zeta_i} = \int_{IR} \left(\frac{\partial P(u_{ji}|\theta, \zeta_i)/\partial \zeta_i}{P(u_{ji}|\zeta_i)} \right) P(u_{j.}|\theta, \zeta) g(\theta|\eta) d\theta,$$

onde a ordem da derivada e da integral em:

$$\frac{\partial P(u_{j.}|\zeta, \eta)}{\partial \zeta_i} = \int_{IR} \left(\frac{\partial}{\partial \zeta_i} \prod_{l=1}^I P(u_{jl}|\theta, \zeta_l) \right) g(\theta|\eta) d\theta$$

pode ser permutada com base no Teorema da Convergência Dominada de Lebesgue (Chow & Teicher, 1978).

Reescrevendo $P(u_{ji}|\theta, \zeta_i)$ como em

$$P(u_{ji} = u_{ji}|\theta, \zeta_i) = P_i^{u_{ji}} Q_i^{1-u_{ji}},$$

tem-se que:

$$\frac{\partial}{\partial \zeta_i} P(u_{ji}|\theta, \zeta_i) = \frac{\partial}{\partial \zeta_i} \left(P_i^{u_{ji}} Q_i^{1-u_{ji}} \right)$$

$$\frac{\partial}{\partial \zeta_i} P(u_{ji}|\theta, \zeta_i) = u_{ji} P_i^{u_{ji}-1} \left(\frac{\partial P_i}{\partial \zeta_i} \right) Q_i^{1-u_{ji}} + (1-u_{ji}) Q_i^{-u_{ji}} \left(-\frac{\partial}{\partial \zeta_i} P_i \right) P_i^{u_{ji}}$$

$$\frac{\partial}{\partial \zeta_i} P(u_{ji}|\theta, \zeta_i) = \left(u_{ji} P_i^{u_{ji}-1} Q_i^{1-u_{ji}} - (1-u_{ji}) Q_i^{-u_{ji}} P_i^{u_{ji}} \right) \frac{\partial P_i}{\partial \zeta_i}.$$

Nota-se que o termo entre parênteses vale 1 quando $u_{ji} = 1$ e vale -1 quando $u_{ji} = 0$, portanto pode-se reescrevê-lo como $(-1)^{u_{ji}+1}$.

Com isso:

$$\frac{\partial}{\partial \zeta_i} P(u_{ji}|\theta, \zeta_i) = (-1)^{u_{ji}+1} \left(\frac{\partial P_i}{\partial \zeta_i} \right).$$

Notando agora que:

$$\frac{(-1)^{u_{ji}+1} P_i Q_i}{P_i^{u_{ji}} Q_i^{1-u_{ji}}} = \begin{cases} Q_i & \text{se } u_{ji} = 1 \\ -P_i & \text{se } u_{ji} = 0, \end{cases}$$

Pode-se reescrever este termo como $u_{ji} - P_i$. Segue que:

$$\frac{\partial P(u_{j.}|\zeta, \eta)}{\partial \zeta_i} = \int_{IR} \left(\frac{\partial P(u_{ji}|\theta, \zeta_i)/\partial \zeta_i}{P(u_{ji}|\zeta_i)} \right) P(u_{j.}|\theta, \zeta) g(\theta|\eta) d\theta,$$

pode ser escrita como:

$$\frac{\partial P(u_{j.}|\zeta, \eta)}{\partial \zeta_i} = \int IR \left[\left(\frac{u_{ji} - P_i}{P_i Q_i} \right) \left(\frac{\partial P_i}{\partial \zeta_i} \right) \right] P(u_{j.}|\theta, \zeta) g(\theta|\eta) d\theta.$$

Por conveniência, considera-se a seguinte ponderação já vista anteriormente:

$$w_i = \frac{P_i^* Q_i^*}{P_i Q_i},$$

onde :

$$P_i^* = \left\{ 1 + e^{-Da_i(\theta - b_i)} \right\}^{-1} \text{ e } Q_i^* = 1 - P_i^*.$$

Com isso, pode-se reescrever a equação

$$\frac{\partial P(u_{j.}|\zeta, \eta)}{\partial \zeta_i} = \int IR \left[\left(\frac{u_{ji} - P_i}{P_i Q_i} \right) \left(\frac{\partial P_i}{\partial \zeta_i} \right) \right] P(u_{j.}|\theta, \zeta) g(\theta|\eta) d\theta$$

como:

$$\frac{\partial P(u_{j.}|\zeta, \eta)}{\partial \zeta_i} = \int IR \left[\left(u_{ji} - P_i \right) \left(\frac{\partial P_i}{\partial \zeta_i} \right) \frac{w_i}{P_i^* Q_i^*} \right] P(u_{j.}|\theta, \zeta) g(\theta|\eta) d\theta.$$

Usando a notação:

$$g_j^*(\theta) \equiv g(\theta|u_{j.}, \zeta, \eta) = \frac{P(u_{j.}|\theta, \zeta) g(\theta|\eta)}{P(u_{j.}|\zeta, \eta)},$$

tem-se que a função de verossimilhança:

$$\frac{\partial \log L(\zeta, \eta)}{\partial \zeta_i} = \sum_{j=1}^s r_j \frac{1}{P(u_{j.}|\zeta, \eta)} \frac{\partial P(u_{j.}|\zeta, \eta)}{\partial \zeta_i},$$

que pode ser escrita como:

$$\frac{\partial \log L(\zeta, \eta)}{\partial \zeta_i} = \sum_{j=1}^s r_j \int IR \left[\left(u_{ji} - P_i \right) \left(\frac{\partial P_i}{\partial \zeta_i} \right) \frac{w_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta.$$

esta agora a obtenção das equações específicas para cada parâmetros do vetor $\zeta_i = (a_i, b_i, c_i)$.

As expressões para as derivadas de P_i são dadas por:

$$\frac{\partial P_{ji}}{\partial a_i} = D(1 - c_i)(\theta_j - b_i)P_{ji}^* Q_{ji}^*,$$

$$\frac{\partial P_{ji}}{\partial b_i} = -Da_i(1 - c_i)P_{ji}^* Q_{ji}^*,$$

$$\frac{\partial P_{ji}}{\partial c_i} = Q_{ji}^*.$$

Com P_{ji} , Q_{ji} , P_{ji}^* e Q_{ji}^* substituídas por P_i , Q_i , P_i^* e Q_i^* , respectivamente.

Para obterem-se as seguintes equações de estimação dos parâmetros:

Para o parâmetro de discriminação (a_i), da expressão

$$\frac{\partial \log L(\zeta, \eta)}{\partial \zeta_i} = \sum_{j=1}^s r_j \int_{IR} \left[(u_{ji} - P_i) \left(\frac{\partial P_i}{\partial \zeta_i} \right) \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta.$$

vem:

$$\begin{aligned} \frac{\partial \log L(\zeta, \eta)}{\partial a_i} &= \sum_{j=1}^s r_j \int_{IR} \left[(u_{ji} - P_i) \left(\frac{\partial P_i}{\partial a_i} \right) \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta \\ \frac{\partial \log L(\zeta, \eta)}{\partial a_i} &= \sum_{j=1}^s r_j \int_{IR} \left[(u_{ji} - P_i) D(1 - c_i) (\theta - b_i) P_i^* Q_i^* \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta \\ \frac{\partial \log L(\zeta, \eta)}{\partial a_i} &= D(1 - c_i) \sum_{j=1}^s r_j \int_{IR} \left[(u_{ji} - P_i) (\theta - b_i) W_i \right] g_j^*(\theta) d\theta. \end{aligned}$$

Para o parâmetro de dificuldade, (b_i), tem-se :

$$\begin{aligned} \frac{\partial \log L(\zeta, \eta)}{\partial b_i} &= \sum_{j=1}^s r_j \int_{IR} \left[(u_{ji} - P_i) \left(\frac{\partial P_i}{\partial b_i} \right) \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta \\ \frac{\partial \log L(\zeta, \eta)}{\partial b_i} &= \sum_{j=1}^s r_j \int_{IR} \left[(u_{ji} - P_i) (-1) D a_i (1 - c_i) P_i^* Q_i^* \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta \\ \frac{\partial \log L(\zeta, \eta)}{\partial b_i} &= -D a_i (1 - c_i) \sum_{j=1}^s r_j \int_{IR} \left[(u_{ji} - P_i) W_i \right] g_j^*(\theta) d\theta. \end{aligned}$$

Para o parâmetro de acerto ao acaso, (c_i), tem-se:

$$\begin{aligned} \frac{\partial \log L(\zeta, \eta)}{\partial c_i} &= \sum_{j=1}^s r_j \int_{IR} \left[(u_{ji} - P_i) \left(\frac{\partial P_i}{\partial c_i} \right) \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta \\ \frac{\partial \log L(\zeta, \eta)}{\partial c_i} &= \sum_{j=1}^s r_j \int_{IR} \left[(u_{ji} - P_i) Q_i^* \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta \\ \frac{\partial \log L(\zeta, \eta)}{\partial c_i} &= \sum_{j=1}^s r_j \int_{IR} \left[(u_{ji} - P_i) \frac{W_i}{P_i^*} \right] g_j^*(\theta) d\theta. \end{aligned}$$

Em resumo, as equações de estimação para os parâmetros a_i , b_i e c_i , para a Máxima Verossimilhança Marginal (MVM) são, respectivamente:

$$\begin{aligned} a_i &: D(1 - c_i) \sum_{j=1}^s r_j \int_{IR} \left[(u_{ji} - P_i) (\theta - b_i) W_i \right] g_j^*(\theta) d\theta = 0, \\ b_i &: -D a_i (1 - c_i) \sum_{j=1}^s r_j \int_{IR} \left[(u_{ji} - P_i) W_i \right] g_j^*(\theta) d\theta = 0, \end{aligned}$$

$$c_i : \sum_{j=1}^s r_j \int_{IR} \left[(u_{ji} - P_i) \frac{W_i}{P_i^*} \right] g_j^*(\theta) d\theta = 0,$$

as quais também não possuem solução explícita, (Andrade 2000, p.52-57).

2.5.4.2 Métodos iterativos

Para aplicação do algoritmo Newton-Raphson, precisa-se das derivadas segundas de $\log L(\zeta, \eta)$. Quando foi desenvolvido as expressões para a estimação de ζ_i na estimação dos parâmetros dos itens, a propriedade de independência local foi suficiente para garantir que os parâmetros dos itens pudessem ser estimados individualmente, pois a derivada segunda de $\log L(\zeta)$ com relação a ζ_i e ζ_l , para $l \neq i$, era nula. Entretanto, na estimação por máxima verossimilhança marginal isso não acontece, levando à necessidade da estimação dos I itens conjuntamente. As expressões para as derivadas segundas são obtidas a partir de:

$$\frac{\partial^2 \log L(\zeta, \eta)}{\partial \zeta_l \partial \zeta_i} = \frac{\partial}{\partial \zeta_l} \left[\frac{\partial \log L(\zeta, \eta)}{\partial \zeta_i} \right]' = \frac{\partial}{\partial \zeta_l} \left[\sum_{j=1}^s r_j \frac{1}{P(u_{j.}|\zeta, \eta)} \frac{\partial P(u_{j.}|\zeta, \eta)}{\partial \zeta_i} \right]'$$

$$\frac{\partial^2 \log L(\zeta, \eta)}{\partial \zeta_l \partial \zeta_i} = \sum_{j=1}^s r_j \left\{ \frac{\partial^2 P(u_{j.}|\zeta, \eta) / (\partial \zeta_l \partial \zeta_i)}{P(u_{j.}|\zeta, \eta)} - \left(\frac{\partial P(u_{j.}|\zeta, \eta) / \partial \zeta_l}{P(u_{j.}|\zeta, \eta)} \right) \left(\frac{\partial P(u_{j.}|\zeta, \eta) / \partial \zeta_i}{P(u_{j.}|\zeta, \eta)} \right) \right\}'.$$

Para $i, l = 1, 2, \dots, I$.

Considerando $\bar{\zeta}^{(t)}$ a estimativa de ζ na iteração t , então na iteração $t + 1$ tem-se que:

$$\bar{\zeta}^{(t+1)} = \bar{\zeta}^{(t)} - [H_{PI}(\bar{\zeta}^{(t)})]^{-1} h_{PI}(\bar{\zeta}^{(t)})$$

Para chegar às expressões de $H_{PI}(\bar{\zeta})$ e $h_{PI}(\zeta_i)$ usadas a seguir, adota-se a notação:

$$v_{ji} = (u_{ji} - P_i) \frac{W_i}{P_i^* Q_i^*} = \frac{(u_{ji} - P_i)}{P_i Q_i}.$$

Segue, de

$$\frac{\partial P(u_{j.}|\zeta, \eta)}{\partial \zeta_i} = \int_{IR} \left(\frac{\partial}{\partial \zeta_i} \prod_{l=1}^I P(u_{jl}|\theta, \zeta_l) \right) g(\theta|\eta) d\theta,$$

$$\frac{\partial P(u_{j.}|\zeta, \eta)}{\partial \zeta_i} = \int_{IR} \left[(u_{ji} - P_i) \left(\frac{\partial P_i}{\partial \zeta_i} \right) \frac{W_i}{P_i^* Q_i^*} \right] P(u_{j.}|\theta, \zeta) g(\theta|\eta) d\theta,$$

que

$$\frac{\partial P(u_{j.}|\theta, \zeta)}{\partial \zeta_i} = \left[v_{ji} \left(\frac{\partial P_i}{\partial \zeta_i} \right) \right] P(u_{j.}|\theta, \zeta)$$

E de

$$\frac{\partial P(u_{j.}|\zeta, \eta)}{\partial \zeta_i} = \int_{IR} \left[\left(\frac{u_{ji} - P_i}{P_i Q_i} \right) \left(\frac{\partial P_i}{\partial \zeta_i} \right) \right] P(u_{j.}|\theta, \zeta) g(\theta|\eta) d\theta,$$

que é a segunda parcela de

$$\frac{\partial^2 \log L(\zeta, \eta)}{\partial \zeta_l \partial \zeta_i'} = \sum_{j=1}^s r_j \left\{ \frac{\partial^2 P(u_{j.}|\zeta, \eta) / (\partial \zeta_l \partial \zeta_i')}{P(u_{j.}|\zeta, \eta)} - \left(\frac{\partial P(u_{j.}|\zeta, \eta) / \partial \zeta_l}{P(u_{j.}|\zeta, \eta)} \right) \left(\frac{\partial P(u_{j.}|\zeta, \eta) / \partial \zeta_i}{P(u_{j.}|\zeta, \eta)} \right)' \right\}$$

é obtida por:

$$h_{i(j)} \equiv \frac{\partial P(u_{j.}|\zeta, \eta) / \partial \zeta_i}{P(u_{j.}|\zeta, \eta)} = \int_{IR} \left[v_{ji} \left(\frac{\partial P_i}{\partial \zeta_i} \right) \right] g_j^*(\theta) d\theta.$$

Com relação à primeira parcela de :

$$\frac{\partial^2 \log L(\zeta, \eta)}{\partial \zeta_l \partial \zeta_i'} = \sum_{j=1}^s r_j \left\{ \frac{\partial^2 P(u_{j.}|\zeta, \eta) / (\partial \zeta_l \partial \zeta_i')}{P(u_{j.}|\zeta, \eta)} - \left(\frac{\partial P(u_{j.}|\zeta, \eta) / \partial \zeta_l}{P(u_{j.}|\zeta, \eta)} \right) \left(\frac{\partial P(u_{j.}|\zeta, \eta) / \partial \zeta_i}{P(u_{j.}|\zeta, \eta)} \right)' \right\},$$

nota-se que

$$\frac{\partial^2 P(u_{j.}|\zeta, \eta)}{\partial \zeta_l \partial \zeta_i'} = \frac{\partial}{\partial \zeta_i} \left\{ \int_{IR} \left[v_{ji} \left(\frac{\partial P_i}{\partial \zeta_i} \right) \right]' P(u_{j.}|\theta, \zeta) g(\theta|\eta) d\theta \right\}.$$

Utilizando $\frac{\partial P(u_{j.}|\theta, \zeta)}{\partial \zeta_i} = \left[v_{ji} \left(\frac{\partial P_i}{\partial \zeta_i} \right) \right] P(u_{j.}|\theta, \zeta)$ e o desenvolvimento na aplicação do

Algoritmo de Newton-Raphson, que resultou em:

$$\frac{\partial v_{ji}}{\partial \zeta_i} = -v_{ji}^2 \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right),$$

tem-se:

$$\begin{aligned} \frac{\partial}{\partial \zeta_i} \left\{ \left[v_{ji} \left(\frac{\partial P_i}{\partial \zeta_i} \right) \right]' P(u_{j.}|\theta, \zeta) \right\} &= \frac{\partial}{\partial \zeta_i} \left[v_{ji} \left(\frac{\partial P_i}{\partial \zeta_i} \right) \right]' P(u_{j.}|\theta, \zeta) + v_{ji} \left(\frac{\partial P(u_{j.}|\theta, \zeta)}{\partial \zeta_i} \right) \left(\frac{\partial P_i}{\partial \zeta_i} \right)' \\ &= \left[-v_{ji}^2 \left(\frac{\partial P_i}{\partial \zeta_i} \right) \left(\frac{\partial P_i}{\partial \zeta_i} \right)' + v_{ji} \left(\frac{\partial^2 P_i}{\partial \zeta_i \partial \zeta_i'} \right) \right] P(u_{j.}|\theta, \zeta) + v_{ji}^2 P(u_{j.}|\theta, \zeta) \left(\frac{\partial P_i}{\partial \zeta_i} \right) \left(\frac{\partial P_i}{\partial \zeta_i} \right)' \\ \frac{\partial}{\partial \zeta_i} \left\{ \left[v_{ji} \left(\frac{\partial P_i}{\partial \zeta_i} \right) \right]' P(u_{j.}|\theta, \zeta) \right\} &= v_{ji} \left(\frac{\partial^2 P_i}{\partial \zeta_i \partial \zeta_i'} \right) P(u_{j.}|\theta, \zeta) \end{aligned}$$

Segue de $\frac{\partial^2 P(u_{j.}|\zeta, \eta)}{\partial \zeta_i \partial \zeta_i'} = \frac{\partial}{\partial \zeta_i} \left\{ \int_{IR} \left[v_{ji} \left(\frac{\partial P_i}{\partial \zeta_i} \right)' \right] P(u_{j.}|\theta, \zeta) g(\theta|\eta) d\theta \right\}$, que

$$\frac{\partial^2 P(u_{j.}|\zeta, \eta)}{\partial \zeta_i \partial \zeta_i'} = \int_{IR} v_{ji} \left(\frac{\partial^2 P_i}{\partial \zeta_i \partial \zeta_i'} \right) P(u_{j.}|\theta, \zeta) g(\theta|\eta) d\theta.$$

Portanto, a primeira parcela em

$$\frac{\partial^2 \log L(\zeta, \eta)}{\partial \zeta_l \partial \zeta_i'} = \sum_{j=1}^s r_j \left\{ \frac{\partial^2 P(u_{j.}|\zeta, \eta) / (\partial \zeta_l \partial \zeta_i')}{P(u_{j.}|\zeta, \eta)} - \left(\frac{\partial P(u_{j.}|\zeta, \eta) / \partial \zeta_l}{P(u_{j.}|\zeta, \eta)} \right) \left(\frac{\partial P(u_{j.}|\zeta, \eta) / \partial \zeta_i}{P(u_{j.}|\zeta, \eta)} \right)' \right\}$$

pode ser escrita como:

$$H_{ii(j)} \equiv \frac{\partial^2 P(u_{j.}|\zeta, \eta) / (\partial \zeta_i \partial \zeta_i')}{P(u_{j.}|\zeta, \eta)} = \int_{IR} v_{ji} \left(\frac{\partial^2 P_i}{\partial \zeta_i \partial \zeta_i'} \right) g_j^*(\theta) d\theta.$$

Por $\frac{\partial P(u_{j.}|\zeta, \eta)}{\partial \zeta_i} = \int_{IR} \left[(u_{ji} - P_i) \left(\frac{\partial P_i}{\partial \zeta_i} \right) \frac{W_i}{P_i^* Q_i^*} \right] P(u_{j.}|\theta, \zeta) g(\theta|\eta) d\theta$, para $l \neq i$

$$\frac{\partial^2 P(u_{j.}|\zeta, \eta)}{\partial \zeta_l \partial \zeta_i'} = \frac{\partial}{\partial \zeta_l} \left\{ \int_{IR} \left[v_{ji} \left(\frac{\partial P_i}{\partial \zeta_i} \right)' \right] P(u_{j.}|\theta, \zeta) g(\theta|\eta) d\theta \right\}$$

$$\frac{\partial^2 P(u_{j.}|\zeta, \eta)}{\partial \zeta_l \partial \zeta_i'} = \int_{IR} v_{ji} \left(\frac{\partial P(u_{j.}|\theta, \zeta)}{\partial \zeta_l} \right) \left(\frac{\partial P_i}{\partial \zeta_i} \right)' g(\theta|\eta) d\theta$$

$$\frac{\partial^2 P(u_{j.}|\zeta, \eta)}{\partial \zeta_l \partial \zeta_i'} = \int_{IR} v_{ji} v_{jl} \left(\frac{\partial P_l}{\partial \zeta_l} \right) \left(\frac{\partial P_i}{\partial \zeta_i} \right)' P(u_{j.}|\theta, \zeta) g(\theta|\eta) d\theta$$

Portanto, para $l \neq i$, a primeira parcela em:

$$\frac{\partial^2 \log L(\zeta, \eta)}{\partial \zeta_l \partial \zeta_i'} = \sum_{j=1}^s r_j \left\{ \frac{\partial^2 P(u_{j.}|\zeta, \eta) / (\partial \zeta_l \partial \zeta_i')}{P(u_{j.}|\zeta, \eta)} - \left(\frac{\partial P(u_{j.}|\zeta, \eta) / \partial \zeta_l}{P(u_{j.}|\zeta, \eta)} \right) \left(\frac{\partial P(u_{j.}|\zeta, \eta) / \partial \zeta_i}{P(u_{j.}|\zeta, \eta)} \right)' \right\},$$

pode ser escrita como:

$$H_{il(j)} \equiv \frac{\partial^2 P(u_{j.}|\zeta, \eta) / (\partial \zeta_l \partial \zeta_i')}{P(u_{j.}|\zeta, \eta)} = \int_{IR} v_{ji} v_{jl} \left(\frac{\partial P_l}{\partial \zeta_l} \right) \left(\frac{\partial P_i}{\partial \zeta_i} \right)' g_j^*(\theta) d\theta.$$

Pode-se agora obter as equações de estimação para ζ . Com as expressões de Newton-Raphson:

$$\frac{\partial P_{ji}^* Q_{ji}^*}{\partial \alpha_i} = (1 - 2P_{ji}^*) \frac{\partial P_{ji}^*}{\partial \alpha_i}, \quad \alpha_i \in \{a_i, b_i, c_i\},$$

$$\frac{\partial^2 P_{ji}}{\partial a_i^2} = D^2(1-c_i) (\theta_j - b_i)^2 P_{ji}^* Q_{ji}^* (1-2P_{ji}^*).$$

$$\frac{\partial^2 P_{ji}}{\partial a_i \partial b_i} = -D(1-c_i) P_{ji}^* Q_{ji}^* \left\{ 1 + Da_i (\theta_j - b_i) (1-2P_{ji}^*) \right\},$$

$$\frac{\partial^2 P_{ji}}{\partial a_i \partial c_i} = -D(\theta_j - b_i) P_{ji}^* Q_{ji}^*,$$

$$\frac{\partial^2 P_{ji}}{\partial b_i^2} = D^2 a_i^2 (1-c_i) P_{ji}^* Q_{ji}^* (1-2P_{ji}^*),$$

$$\frac{\partial^2 P_{ji}}{\partial b_i \partial c_i} = Da_i P_{ji}^* Q_{ji}^*,$$

$$\frac{\partial^2 P_{ji}}{\partial c_i^2} = \frac{\partial Q_{ji}^*}{\partial c_i} = 0.$$

De $\left(\frac{\partial^2 P_i}{\partial \zeta_i \partial \zeta_i'} \right)$, $i = 1, 2, \dots, I$. onde:

$$h_{PI}(\zeta) = \begin{pmatrix} h(\zeta_1) \\ \vdots \\ h(\zeta_I) \end{pmatrix} \quad \text{e} \quad H_{PI}(\zeta) = \begin{pmatrix} H(\zeta_1, \zeta_1) & \dots & H(\zeta_1, \zeta_I) \\ \vdots & \ddots & \vdots \\ H(\zeta_I, \zeta_1) & \dots & H(\zeta_I, \zeta_I) \end{pmatrix},$$

Sejam:

$$h_i = (P_i^* Q_i^*)^{-1} \left(\frac{\partial P_i}{\partial \zeta_i} \right) = \begin{pmatrix} D(1-c_i)(\theta - b_i) \\ -Da_i(1-c_i) \\ \frac{1}{P_i^*} \end{pmatrix},$$

$$H_{ii} = (P_i^* Q_i^*)^{-1} \left(\frac{\partial^2 P_i}{\partial \zeta_i \partial \zeta_i'} \right) = \begin{pmatrix} D^2(1-c_i)(\theta - b_i)^2(1-2P_i^*) & \cdot & \cdot \\ -D(1-c_i) \left\{ 1 + Da_i(\theta - b_i)(1-2P_i^*) \right\} & D^2 a_i^2 (1-c_i)(1-2P_i^*) & \cdot \\ -D(\theta - b_i) & Da_i & 0 \end{pmatrix}$$

e, para $i \neq l$,

$$H_{il} = h_i h_l' = (P_i^* Q_i^*)^{-1} (P_l^* Q_l^*)^{-1} \left(\frac{\partial P_i}{\partial \zeta_i} \right) \left(\frac{\partial P_l}{\partial \zeta_l} \right)',$$

$$H_{il} = \begin{pmatrix} D^2(1-c_i)(\theta-c_l)(\theta-b_i)(\theta-b_l) & -D^2a_l(1-c_i)(1-c_l)(\theta-b_i) & D(1-c_i)(\theta-b_i)/P_l^* \\ -D^2a_i(1-c_i)(1-c_l)(\theta-b_l) & D^2a_i a_l(1-c_i)(1-c_l) & -Da_i(1-c_i)/P_l^* \\ -D(1-c_l)(\theta-b_l)/P_i^* & -Da_l(1-c_l)/P_i^* & [P_i^* P_l^*]^{-1} \end{pmatrix}.$$

Retornando a $H_{ii(j)} = \int_{IR} v_{ji} \left(\frac{\partial^2 P_i}{\partial \zeta_i \partial \zeta_i'} \right) g_j^*(\theta) d\theta$, tem-se a primeira parcela de

$$\frac{\partial^2 \log L(\zeta, \eta)}{\partial \zeta_l \partial \zeta_i'} = \sum_{j=1}^s r_j \left\{ \frac{\partial^2 P(u_j | \zeta, \eta) / (\partial \zeta_l \partial \zeta_i')}{P(u_j | \zeta, \eta)} - \left(\frac{\partial P(u_j | \zeta, \eta) / \partial \zeta_l}{P(u_j | \zeta, \eta)} \right) \left(\frac{\partial P(u_j | \zeta, \eta) / \partial \zeta_i}{P(u_j | \zeta, \eta)} \right)' \right\}$$

pode ser reescrita como: $H_{ii(j)} = \int_R (u_{ji} - P_i) w_i H_{ii} g_j^*(\theta) d\theta$ e, para $i \neq l$,

$$H_{il(j)} = \int_R (u_{ji} - P_i)(u_{jl} - P_l) w_i H_{il} g_j^*(\theta) d\theta$$

Com isso, chega-se a

$$H(\zeta_i, \zeta_l) = \frac{\partial^2 \log L(\zeta, \eta)}{\partial \zeta_l \partial \zeta_i'} = \sum_{j=1}^s r_j \{ H_{il(j)} - h_{i(j)} h_{l(j)}' \}$$

Para aplicar o algoritmo “Scoring” de Fisher, nota-se

$$E[H_{il(j)}] = 0, \quad i, l = 1, 2, \dots, I \text{ e } j = 1, 2, \dots, n.$$

Segue então que :

$$\Delta(\zeta_i, \zeta_l) = E[H(\zeta_i, \zeta_l)] = - \sum_{j=1}^s r_j [h_{i(j)} h_{l(j)}'] \quad (\text{Andrade 2000, p.58-59}).$$

2.5.4.3 Métodos de Quadratura Gaussiana

Segundo (Andrade 2000, p.59), um problema importante encontrado na implementação da estimação dos parâmetros dos itens. Notando que as equações:

$$a_i : D(1-c_i) \sum_{j=1}^s r_j \int_{IR} (u_{ji} - P_i)(\theta - b_i) w_i g_j^*(\theta) d\theta = 0,$$

$$b_i : -Da_i(1-c_i) \sum_{j=1}^s r_j \int_{IR} (u_{ji} - P_i) w_i g_j^*(\theta) d\theta = 0,$$

$$c_i : \sum_{j=1}^s r_j \int_{IR} (u_{ji} - P_i) \frac{w_i}{P_i^*} g_j^*(\theta) d\theta = 0,$$

envolvem integrais que não apresentam solução analítica. Por conta disso, algum meio deve ser encontrado para a solução (aproximação) numérica de uma integral. Embora existam muitos

métodos de aproximações de integrais, na TRI têm sido freqüente a aplicação do método Hermite-Gauss, usualmente denominado de método de quadratura gaussiana. Se $g(\theta|\eta)$ é uma função contínua com integral finita, ela pode ser aproximada, para qualquer grau de precisão, por uma outra função que assume um número finito de pontos. Dessa forma, o problema de obter a integral de uma função contínua é substituído pela obtenção da soma das áreas de um número finito, diga-se q , de retângulos. Os pontos médios de cada retângulo, $\bar{\theta}_k$, $k = 1, 2, \dots, q$, são denominados de nós (ou pontos de quadratura). Cada nó tem um peso $A_k \equiv A(\bar{\theta}_k)$ associado que leva em conta a altura $g(\bar{\theta}_k|\eta)$ e a largura (Δ_k) do respectivo intervalo, tal como $A_k = g(\bar{\theta}_k|\eta) \times \Delta_k$. Os valores $\bar{\theta}_k$ e A_k são obtidos resolvendo-se um conjunto de equações que envolvem a função $g(\theta|\eta)$ e o número de nós. Uma tabela para $\bar{\theta}_k$ e A_k relativa a função gaussiana pode ser encontrada em Stroud & Secrest (1966). Para adaptar essa tabela para o caso em que $g(\bar{\theta}_k|\eta)$ representa a fdp de uma variável $N(0, 1)$, basta multiplicar os nós $\bar{\theta}_k$ por $\sqrt{2}$ e dividir os pesos A_k por $\sqrt{\pi}$, (Andrade 2000, p.59).

2.5.4.3.1 Equações de Estimação Conjunta em Forma de Quadratura

Conhecidos os nós $\bar{\theta}_k$ e os pesos, A_k , $k = 1, 2, \dots, q$, com $A_k = g(\bar{\theta}_k|\eta) \times \Delta_k$.

Com isso, pode-se escrever:

$$P(u_{j.}|\bar{\theta}_k, \zeta) = \prod_{i=1}^I \left[P_{ki}^{u_{ji}} Q_{ki}^{1-u_{ji}} \right]$$

$$P(u_{j.}|\bar{\theta}_k, \zeta) g(\bar{\theta}_k|\eta) = P(u_{j.}|\bar{\theta}_k, \zeta) A_k \Delta_k^{-1}$$

$$P(u_{j.}|\zeta, \eta) \equiv \sum_{k=1}^q P(u_{j.}|\bar{\theta}_k, \zeta) g(\bar{\theta}_k|\eta) \Delta_k = \sum_{k=1}^q P(u_{j.}|\bar{\theta}_k, \zeta) A_k.$$

Segue que

$$g_j^*(\theta) \equiv g(\theta|u_{j.}, \zeta, \eta) = \frac{P(u_{j.}|\theta, \zeta) g(\theta|\eta)}{P(u_{j.}|\zeta, \eta)},$$

pode ser escrita, em forma de quadratura, como:

$$g_j^*(\bar{\theta}_k) \equiv \frac{P(u_{j.} | \bar{\theta}_k, \zeta) A_k}{\sum_{k=1}^q P(u_{j.} | \bar{\theta}_k, \zeta) A_k} \Delta_k^{-1}.$$

Por exemplo, voltando à função de verossimilhança para a_i dada por

$$a_i : D(1-c_i) \sum_{j=1}^S r_j \int_{IR} \left[(u_{ji} - P_i)(\theta - b_i) W_i \right] g_j^*(\theta) d\theta = 0,$$

pode-se reescrevê-la em forma de quadratura como:

$$\begin{aligned} \frac{\partial \log L(\zeta, \eta)}{\partial a_i} &= D(1-c_i) \sum_{j=1}^S r_j \int_{IR} \left[(u_{ji} - P_i)(\theta - b_i) W_i \right] g_j^*(\theta) d\theta \\ \frac{\partial \log L(\zeta, \eta)}{\partial a_i} &\equiv D(1-c_i) \sum_{j=1}^S \sum_{k=1}^q r_j \left[(u_{ji} - P_{ki})(\bar{\theta}_k - b_i) W_{ki} \right] g_j^*(\bar{\theta}_k) \Delta_k \end{aligned}$$

Para que a expressão em forma de quadratura fique o mais parecida possível com a original, pode-se redefinir a quantidade

$$g_j^*(\theta_k) \text{ de } g_j^*(\bar{\theta}_k) \equiv \frac{P(u_{j.} | \bar{\theta}_k, \zeta) A_k}{\sum_{k=1}^q P(u_{j.} | \bar{\theta}_k, \zeta) A_k} \Delta_k^{-1},$$

por

$$g_j^*(\theta_k) = \frac{P(u_{j.} | \bar{\theta}_k, \zeta) A_k}{\sum_{k=1}^q P(u_{j.} | \bar{\theta}_k, \zeta) A_k}.$$

Desta forma, a função de verossimilhança para a_i fica:

$$a_i : D(1-c_i) \sum_{j=1}^S \sum_{k=1}^q r_j \left[(u_{ji} - P_{ki})(\bar{\theta}_k - b_i) W_{ki} \right] g_j^*(\bar{\theta}_k).$$

De forma análoga, tem-se as equações de estimação em forma de quadratura para os parâmetros b_i e c_i são, respectivamente:

$$\begin{aligned} b_i : -D a_i (1-c_i) \sum_{j=1}^S \sum_{k=1}^q r_j \left[(u_{ji} - P_{ki}) W_{ki} \right] g_j^*(\bar{\theta}_k) &= 0, \\ c_i : \sum_{j=1}^S \sum_{k=1}^q r_j \left[(u_{ji} - P_{ki}) \frac{W_{ki}}{P_{ki}} \right] g_j^*(\bar{\theta}_k) &= 0. \end{aligned}$$

Sendo ressaltado que a função deve ser calculada por

$$g_j^*(\theta_k) = \frac{P(u_{j.} | \bar{\theta}_k, \zeta) A_k}{\sum_{k=1}^q P(u_{j.} | \bar{\theta}_k, \zeta) A_k},$$

nas equações de estimação dos parâmetros:

$$\begin{aligned}
a_i &: D(1-c_i) \sum_{j=1}^S \sum_{k=1}^q r_j \left[(u_{ji} - P_{ki})(\bar{\theta}_k - b_i) w_{ki} \right] g_j^*(\bar{\theta}_k), \\
b_i &: -Da_i(1-c_i) \sum_{j=1}^S \sum_{k=1}^q r_j \left[(u_{ji} - P_{ki}) w_{ki} \right] g_j^*(\bar{\theta}_k) = 0, \\
c_i &: \sum_{j=1}^S \sum_{k=1}^q r_j \left[(u_{ji} - P_{ki}) \frac{w_{ki}}{P_{ki}} \right] g_j^*(\bar{\theta}_k) = 0.
\end{aligned}$$

Novamente, estas equações não apresentam soluções explícitas para os EMV dos parâmetros dos itens. Para aplicação dos procedimentos Newton-Raphson ou "Scoring" de Fisher deve-se notar que as derivadas segundas de $\log L(\zeta, \eta)$ com relação a ζ_i e ζ_l , para $i \neq l$, não são nulas, o que leva à necessidade da estimação dos parâmetros dos I itens simultaneamente. Isso pode gerar uma grande limitação na estimação de um número alto de itens devido à necessidade da inversão de matrizes de dimensões $3I \times 3I$. A proposta de Bock & Aitkin (1981), apresentada a seguir, contorna este problema, (Andrade 2000, p.60-61).

2.5.4.4 Abordagem de Bock & Aitkin

Uma reformulação da abordagem de Bock & Lieberman, que foi considerada satisfatória do ponto de vista computacional, foi proposta por Bock & Aitkin (1981). Esta reformulação teve como base a suposição de que os itens são independentes, de forma que:

$$\frac{\partial^2 \log L(\zeta, \eta)}{\partial \zeta_i \partial \zeta_l} = 0 \quad \text{para } i \neq l.$$

Esta suposição modifica a matriz $H_{PI}(\zeta)$ tornando-a bloco diagonal, uma situação similar à da Estimação Conjunta: Itens e Proficiências, onde eram estimados os parâmetros dos itens e as proficiências conjuntamente. Naquele caso, a independência local foi suficiente para garantir

$\frac{\partial^2 \log L(\zeta, \eta)}{\partial \zeta_i \partial \zeta_l} = 0$ e, assim, possibilitar que os itens fossem estimados individualmente, fixadas as

proficiências. A proposta de Bock & Aitkin foi adotar a independência entre os itens de forma a possibilitar que os itens sejam estimados individualmente. Vale notar que as suposições de independência local e a suposição de independência dos itens são completamente diferentes. A primeira está relacionada às respostas dos indivíduos, enquanto a segunda se refere apenas aos

ítems.

Com esta construção, a estimação pode ser feita adotando as mesmas expressões desenvolvidas na seção anterior, fazendo a adaptação devida a $\frac{\partial^2 \log L(\zeta, \eta)}{\partial \zeta_i \partial \zeta_l} = 0$. Entretanto,

Bock & Aitkin sugerem que a obtenção das estimativas de máxima verossimilhança seja feita através do algoritmo EM, introduzido por Dempster, Laird & Rubin (1977), e por isso algumas alterações nas expressões da seção anterior serão necessárias.

De,

$$a_i : D(1-c_i) \sum_{j=1}^S r_j \int_{IR} \left[(u_{ji} - P_i)(\theta - b_i) w_i \right] g_j^*(\theta) d\theta = 0,$$

tem-se:

$$\begin{aligned} \frac{\partial \log L(\zeta, \eta)}{\partial a_i} &= D(1-c_i) \sum_{j=1}^S r_j \int_{IR} \left[(u_{ji} - P_i)(\theta - b_i) w_i \right] g_j^*(\theta) d\theta \\ \frac{\partial \log L(\zeta, \eta)}{\partial a_i} &= D(1-c_i) \sum_{j=1}^S r_j \int_{IR} (\theta - b_i) \left[u_{ji} g_j^*(\theta) - P_i g_j^*(\theta) \right] w_i d\theta \\ \frac{\partial \log L(\zeta, \eta)}{\partial a_i} &= D(1-c_i) \int_{IR} (\theta - b_i) \left[\sum_{j=1}^S r_j u_{ji} g_j^*(\theta) - P_i \sum_{j=1}^S r_j g_j^*(\theta) \right] w_i d\theta \\ \frac{\partial \log L(\zeta, \eta)}{\partial a_i} &= D(1-c_i) \int_{IR} (\theta - b_i) \left[r_i(\theta) - P_i f_i(\theta) \right] w_i d\theta \end{aligned}$$

onde :

$$r_i(\theta) = \sum_{j=1}^S r_j u_{ji} g_j^*(\theta), \quad f_i(\theta) = \sum_{j=1}^S r_j g_j^*(\theta).$$

Lembrando que $g_j^*(\theta)$ é a distribuição condicional de θ_j dado u_j , então $f_i(\theta)$ representa o número esperado de indivíduos, dentre os que responderam o item i em uma população de tamanho n , que têm proficiência θ . Para a quantidade $r_i(\theta)$ contribuem apenas os indivíduos que responderam corretamente ao item i . Logo, esta quantidade representa o número esperado de indivíduos, dentre os que responderam corretamente ao item i em uma população de tamanho n , que têm proficiência θ .

Analogamente, de

$$b_i : -D a_i (1-c_i) \sum_{j=1}^S r_j \int_{IR} \left[(u_{ji} - P_i) w_i \right] g_j^*(\theta) d\theta = 0,$$

$$c_i : \sum_{j=1}^S r_j \int_{IR} \left[(u_{ji} - P_i) \frac{W_i}{P_i^*} \right] g_j^*(\theta) d\theta = 0,$$

tem-se :

$$\frac{\partial \log L(\zeta, \eta)}{\partial b_i} = -Da_i(1 - c_i) \int_{IR} [r_i(\theta) - P_i f_i(\theta)] W_i d\theta,$$

$$\frac{\partial \log L(\zeta, \eta)}{\partial c_i} = \int_{IR} [r_i(\theta) - P_i f_i(\theta)] W_i d\theta, \text{ (Andrade 2000, p.62-63).}$$

2.5.4.4.1 Equações de Estimação em Forma de Quadratura

Considerando conhecidos os nós $\bar{\theta}_k$ e os pesos, A_k , $k = 1, 2, \dots, q$, têm-se as equações de estimação em forma de quadratura para os parâmetros a_i , b_i e c_i são, respectivamente:

$$a_i : D(1 - c_i) \sum_{k=1}^q (\bar{\theta}_k - b_i) [r_{ki} - P_{ki} f_{ki}] W_{ki} = 0,$$

$$b_i : -Da_i(1 - c_i) \sum_{k=1}^q [r_{ki} - P_{ki} f_{ki}] W_{ki} = 0,$$

$$c_i : \sum_{k=1}^q [r_{ki} - P_{ki} f_{ki}] \frac{W_{ki}}{P_{ki}^*} = 0,$$

onde:

$$r_{ki} = \sum_{j=1}^S r_j u_{ji} g_{jk}^*, \quad f_{ki} = \sum_{j=1}^S r_j g_{jk}^* \quad e \quad g_{jk}^* = g_j^*(\bar{\theta}_k), \text{ (Andrade 2000, p.63).}$$

2.5.4.5 Aplicação do algoritmo EM.

O algoritmo EM é um processo iterativo para determinação de estimativas de máxima verossimilhança de parâmetros de modelos de probabilidade na presença de variáveis aleatórias não observadas. Cada iteração deste processo é feita em dois passos: Esperança (E) e Maximização (M). No caso da TRI, o objetivo é obter estimativas de ζ na presença das variáveis não observadas θ . Neste caso, $u_{..}$ representa o vetor de dados incompletos e $(u_{..}, \theta)$ o vetor de dados completos. Seja $f(u_{..}, \theta | \zeta)$ a densidade conjunta dos dados completos. Se $\hat{\zeta}^{(k)}$ é uma estimativa de ζ na iteração t , então os passos EM para obtenção de $\hat{\zeta}^{(k+1)}$ são:

$$\text{Passo E: Calcular } E \left[\log f(u_{..}, \theta | \zeta) \middle| u_{..}, \hat{\zeta}^{(k)} \right],$$

Passo M: Obter $\hat{\zeta}^{(k+1)}$, que maximiza a função do Passo E.

No passo M a maximização pode ser feita pelo algoritmo Newton-Raphson ou "Scoring" de Fisher. Com a suposição de que os itens são independentes, $\frac{\partial^2 \log L(\zeta, \eta)}{\partial \zeta_i \partial \zeta_l} = 0$, a matriz de derivadas segundas torna-se bloco diagonal, possibilitando que os parâmetros dos itens sejam estimados individualmente, eliminando o problema de trabalhar com matrizes de ordem $3I \times 3I$ e passando a operar com matrizes 3×3 .

Há três formas do algoritmo EM, distinguidas pela relação entre a função (densidade) de probabilidade e a forma da família exponencial. A primeira forma se aplica quando a função é um membro regular da família exponencial; a segunda, quando a função não é um membro regular da família exponencial, mas um membro da família exponencial curvada (formada por distribuições em que há restrições no espaço paramétrico) e a terceira, quando a função não tem nenhuma relação com a família exponencial.

Se a Função Resposta ao Item é um membro regular da família exponencial, o procedimento torna-se relativamente simples, Embora o modelo logístico de um parâmetro-ML1 (modelo de Rasch) seja membro da família exponencial, os modelos de 2 e 3 parâmetros não são. Portanto, a terceira forma do algoritmo EM deve ser aplicada nestes casos.

Para descrever brevemente o algoritmo EM aplicado à TRI, supondo que as proficiências estão restritas a um conjunto de q valores, $\bar{\theta}_k$, $k = 1, 2, \dots, q$, com probabilidades π_k , $k = 1, 2, \dots, q$. (Essa suposição pode ser feita porque as aproximações de integrais são feitas por métodos de quadratura, e os valores $\bar{\theta}_k$ corresponderão aos pontos de quadratura). Seja f_{ki} o número de indivíduos com proficiência $\bar{\theta}_k$ respondendo ao item i , $f_i = (f_{1i}, f_{2i}, \dots, f_{qi})'$, com $\sum_{k=1}^q f_{ki} = N$, $f = \begin{pmatrix} f_1 & f_2 & \dots & f_I \end{pmatrix}$. Similarmente, seja r_{ki} os números de indivíduos com proficiência $\bar{\theta}_k$ respondendo corretamente ao item i , $r_i = (r_{1i}, r_{2i}, \dots, r_{qi})'$ e $r = (r_1, r_2, \dots, r_I)$.

Estas definições se assemelham bastante quando se trata da estimação dos parâmetros dos itens com as proficiências conhecidas e agrupadas em q categorias. Vê-se que, de fato, os resultados são muito similares. Entretanto, na Estimação dos Parâmetros dos Itens, as frequências f_{ki} e r_{ki} eram conhecidas, e no caso atual estas quantidades são desconhecidas. Essa é a grande vantagem do algoritmo EM, onde f_{ki} e r_{ki} podem ser tratadas como quantidades não observadas.

Se os n indivíduos que responderão ao item i são selecionados ao acaso da população, a

probabilidade conjunta que os f_{ki} indivíduos tenham proficiências $\bar{\theta}_k$, $k = 1, \dots, q$, é dada pela distribuição multinomial:

$$P(f_i|\pi) = \frac{n!}{\prod_{k=1}^q f_{ki}!} \prod_{k=1}^q \pi_j^{f_{ki}}, \quad i = 1, 2, \dots, I.$$

Dados f_{ki} e $\bar{\theta}_k$, a probabilidade de ocorrerem r_{ki} acertos ao item i dentre as f_{ki} tentativas por indivíduos com proficiência $\bar{\theta}_k$ é:

$$P(r_{ki}|f_{ki}, \bar{\theta}_k) = \binom{f_{ki}}{r_{ki}} P_{ki}^{r_{ki}} Q_{ki}^{f_{ki}-r_{ki}},$$

onde P_{ki} é a Função de Resposta ao Item adotada com θ_j substituída por $\bar{\theta}_k$. A probabilidade conjunta de f e r , dados $\bar{\theta} = (\bar{\theta}_1, \bar{\theta}_2, \dots, \bar{\theta}_q)$ e π , é:

$$\begin{aligned} P(f, r|\bar{\theta}, \pi) &= P(f|\bar{\theta}, \pi) P(r|f, \bar{\theta}, \pi) = P(f|\pi) P(r|h, \bar{\theta}) \\ P(f, r|\bar{\theta}, \pi) &= \left\{ \prod_{i=1}^I P(f_i|\pi) \right\} \left\{ \prod_{i=1}^I \prod_{k=1}^q P(r_{ki}|f_{ki}, \bar{\theta}_k) \right\} \\ \log L(\zeta) &= \log P(f|\pi) + \sum_{i=1}^I \sum_{k=1}^q \log P(r_{ki}|f_{ki}, \bar{\theta}_k) \\ \log L(\zeta) &= \log P(f|\pi) + \sum_{i=1}^I \sum_{k=1}^q \left\{ \log \binom{f_{ki}}{r_{ki}} + r_{ki} \log P_{ki} + (f_{ki} - r_{ki}) \log Q_{ki} \right\} \\ \log L(\zeta) &= C + \sum_{k=1}^q \sum_{i=1}^I \left\{ r_{ki} \log P_{ki} + (f_{ki} - r_{ki}) \log Q_{ki} \right\}, \text{ onde:} \\ C &= \log P(f|\pi) + \sum_{i=1}^I \sum_{k=1}^q \log \binom{f_{ki}}{r_{ki}} \end{aligned}$$

é constante com relação a ζ . Tem-se (f, r) são não-observáveis, mas tomando a esperança da log-verossimilhança, condicional em $u_{..}$ e ζ , e usando a notação:

$$\bar{r}_{ki} = E[r_{ki}|u_{..}, \zeta], \quad \bar{f}_{ki} = E[f_{ki}|u_{..}, \zeta] \quad \text{e} \quad \bar{C} = E[C|u_{..}, \zeta],$$

obtem-se,

$$E[\log L(\zeta)] = \bar{C} + \sum_{i=1}^I \sum_{k=1}^q \left\{ \bar{r}_{ki} \log P_{ki} + (\bar{f}_{ki} - \bar{r}_{ki}) \log Q_{ki} \right\}$$

Nota-se que esta expressão equivale a

$$\log L(\zeta) = C + \sum_{k=1}^q \sum_{i=1}^I \left\{ r_{ki} \log P_{ki} + (f_{ki} - r_{ki}) \log Q_{ki} \right\}.$$

As primeiras parcelas nessas duas expressões são constantes com relação a ζ . Os termos restantes são, praticamente, os mesmos, com f_{ki} e r_{ki} substituídos por \bar{f}_{ki} e \bar{r}_{ki} , respectivamente. Portanto, maximizar a equação:

$$E[\log L(\zeta)] = \bar{C} + \sum_{i=1}^I \sum_{k=1}^q \{ \bar{r}_{ki} \log P_{ki} + (\bar{f}_{ki} - \bar{r}_{ki}) \log Q_{ki} \},$$

com relação a ζ_i é equivalente a maximizar

$$\log L(\zeta) = C + \sum_{k=1}^q \sum_{i=1}^I \left\{ r_{ki} \log P_{ki} + (f_{ki} - r_{ki}) \log Q_{ki} \right\},$$

e representa o Passo E do algoritmo EM. Mais especificamente, os passos E e M são:

Passo E : Usar os pontos de quadratura $\bar{\theta}_k$, os pesos A_k , $k = 1, 2, \dots, q$ e estimativas iniciais dos parâmetros dos itens, $\bar{\zeta}_i$, $i = 1, 2, \dots, I$, para gerar $g_j^*(\bar{\theta}_k)$ e, posteriormente, \bar{r}_{ki} e \bar{f}_{ki} , $i = 1, 2, \dots, I$ e $k = 1, 2, \dots, q$.

Passo M : Com ‘r’ e ‘f’ obtidos no Passo E, resolver as equações de estimação para ζ_i , $i = 1, 2, \dots, I$, usando o algoritmo Newton-Raphson ou “Scoring” de Fisher através das expressões da Estimação dos Parâmetros dos Itens.

Estes passos compõem cada iteração do algoritmo EM, as quais serão repetidas até que algum critério de parada seja alcançado. Após a finalização do processo, os erros-padrão são obtidos com o uso de

$$I(\zeta_i) = -E \left(\frac{\partial^2 \log L(\zeta)}{\partial \zeta_i \partial \zeta_i} \right) = -\Delta(\zeta_i), \text{ (Andrade 2000, p.64-66).}$$

2.5.5 Estimação Bayesiana

A estimação por máxima verossimilhança apresenta problemas na estimação de itens que são respondidos corretamente, ou incorretamente, por todos os indivíduos, e também das proficiências de indivíduos que responderam corretamente, ou incorretamente, a todos os itens. Além disso, há a possibilidade de que as estimativas dos parâmetros dos itens caiam fora do intervalo esperado, tal como valores de a_i negativos, ou valores de c_i fora do intervalo [0,1]. A metodologia bayesiana apresenta uma solução em que estes problemas são contornados.

Há várias propostas para a estimação bayesiana dos parâmetros de interesse na TRI. A

mais utilizada é a Estimac o Bayesiana Marginal proposta por Mislevy (1986), que   uma generalizac o da proposta de Bock & Aitkin (1981). Basicamente, a estimac o bayesiana consiste em estabelecer distribuic es a priori para os par metros de interesse, construir uma nova fun o denominada distribuic o a posteriori e estimar os par metros de interesse com base em alguma caracter stica dessa distribuic o. Considera-se que as componentes de ζ s o vari veis aleat rias independentes e cont nuas, com distribuic es especificadas. Por tratar-se de uma extens o da proposta de Bock & Aitkin, a estimac o   feita por m xima verossimilhan a marginal, em duas etapas.

Para tornar o tratamento mais geral, considera-se que a distribuic o da habilidade/profici ncia   fun o de um vetor de par metros η , com densidade $g(\theta|\eta)$, e que a distribuic o de ζ_i $i = 1, 2, \dots, I$,   fun o de um vetor de par metros τ , com densidade $g(\zeta|\tau)$. Pode-se, ainda, estabelecer distribuic es a priori para os par metros τ e η , diga-se $f(\tau)$ e $g(\eta)$. Para definir a m trica, diga-se (0,1), em que os par metros dos itens (e, posteriormente, as profici ncias) s o estimados, adota-se uma distribuic o degenerada para η em (0,1) ou uma distribuic o que tenha vetor de m dias (0,1) e vari ncias muito pequenas. A primeira op o equivale a eliminar a fun o $g(\eta)$, mas para tornar o tratamento mais geral procura-se mant -la no desenvolvimento da teoria. Com isso, a densidade conjunta desses par metros  :

$$f(\theta, \zeta, \eta, \tau) = f(\zeta|\tau)g(\theta|\eta)f(\tau)g(\eta) = \left\{ \prod_{i=1}^I f(\zeta_i|\tau) \right\} \left\{ \prod_{j=1}^n g(\theta_j|\eta) \right\} f(\tau)g(\eta).$$

Para fazer infer ncias sobre todos esses par metros, deve-se basear na distribuic o a posteriori:

$$f(\theta, \zeta, \eta, \tau|u_{..}) \propto L(u_{..}; \theta, \zeta) f(\zeta|\tau) g(\theta|\eta) f(\tau) g(\eta).$$

Entretanto, geralmente o interessado   em um n mero reduzido de par metros. Nesse caso, deve-se trabalhar com uma posteriori que seja fun o apenas dos par metros de interesse.

2.5.5.1 Estimac o dos Par metros dos Itens utilizando a Metodologia Bayesiana

Para fazer infer ncias com rela o aos par metros dos itens,   conveniente “marginalizar” a posteriori integrando com rela o a θ e τ , obtendo a distribuic o a posteriori de ζ e η :

$$f^*(\zeta, \eta|u_{..}) = C \iint P(u_{..}; \theta, \zeta) f(\zeta|\tau) g(\theta|\eta) f(\tau) g(\eta) d\theta d\tau$$

$$= Cg(\eta) \left\{ \int f(\zeta|\tau)f(\tau)d\tau \right\} \left\{ \int P(u_{..};\theta,\zeta)g(\theta|\eta)d\theta \right\} \\ \propto L(\zeta,\eta)f(\zeta)g(\eta),$$

onde C representa uma constante,

$$L(\zeta,\eta) \equiv P(u_{..};\zeta,\eta) \text{ e } f(\zeta) = \int f(\zeta|\tau)f(\tau)d\tau.$$

Como estimador de ζ pode-se escolher alguma característica de $f^*(\zeta, \eta|u_{..})$, sendo que as mais adotadas são a média e a moda. Considera-se a moda à posteriori como o estimador de ζ , ou seja, o valor de ζ que maximiza a posteriori marginal. Tem-se que:

$$\log f^*(\zeta, \eta|u_{..}) = \text{Const} + \log L(\zeta, \eta) + \log f(\zeta) + \log g(\eta),$$

onde o primeiro termo representa uma constante. Pela suposição de independência entre os itens, a estimação será feita um item por vez. Notando que a última parcela não é função de ζ_i , tem-se que as equações de estimação para os parâmetros dos itens ζ_i , $i = 1, 2, \dots, I$, são dadas por :

$$\frac{\partial f^*(\zeta, \eta|u_{..})}{\partial \zeta_i} = \frac{\partial \log L(\zeta, \eta)}{\partial \zeta_i} + \frac{\partial \log f(\zeta)}{\partial \zeta_i} = 0.$$

$$\text{A primeira parcela de } \frac{\partial f^*(\zeta, \eta|u_{..})}{\partial \zeta_i} = \frac{\partial \log L(\zeta, \eta)}{\partial \zeta_i} + \frac{\partial \log f(\zeta)}{\partial \zeta_i} = 0,$$

é exatamente a mesma obtida em

$$\frac{\partial \log L(\zeta, \eta)}{\partial \zeta_i} = \sum_{j=1}^S r_j \int_{IR} \left[(u_{ji} - P_i) \left(\frac{\partial P_i}{\partial \zeta_i} \right) \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta.$$

A abordagem bayesiana adiciona uma nova parcela

$$\frac{\partial \log L(\zeta, \eta)}{\partial \zeta_i} = \sum_{j=1}^S r_j \int_{IR} \left[(u_{ji} - P_i) \left(\frac{\partial P_i}{\partial \zeta_i} \right) \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta$$

relativa à distribuição a priori associada aos parâmetros dos itens.

A primeira parcela de

$$\frac{\partial f^*(\zeta, \eta|u_{..})}{\partial \zeta_i} = \frac{\partial \log L(\zeta, \eta)}{\partial \zeta_i} + \frac{\partial \log f(\zeta)}{\partial \zeta_i} = 0.$$

relativa às componentes de ζ_I é dada por:

$$a_i : D(1 - c_i) \sum_{j=1}^S r_j \int_{IR} \left[(u_{ji} - P_i) (\theta - b_i) \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta = 0,$$

$$b_i : -Da_i(1 - c_i) \sum_{j=1}^S r_j \int_{IR} \left[(u_{ji} - P_i) \frac{W_i}{P_i^* Q_i^*} \right] g_j^*(\theta) d\theta = 0,$$

$$c_i : \sum_{j=1}^s r_j \int_{IR} \left[(u_{ji} - P_i) \frac{W_i}{P_i^*} \right] g_j^*(\theta) d\theta = 0,$$

A segunda parcela de

$$\frac{\partial f^*(\zeta, \eta | u_{..})}{\partial \zeta_i} = \frac{\partial \log L(\zeta, \eta)}{\partial \zeta_i} + \frac{\partial \log f(\zeta)}{\partial \zeta_i} = 0,$$

depende da priori adotada para cada parâmetro. Como o parâmetro a_i deve ser positivo, b_i pode assumir qualquer valor real e c_i deve estar no intervalo $[0, 1]$, devem-se assumir distribuições que levam em conta essas limitações e isso exige um tratamento diferenciado para cada um destes parâmetros

2.5.5.1.1 Distribuição a priori para a_i

Geralmente, adotam-se as distribuições Log-normal ou Chi-Quadrado para a_i . Nesta definição, supõe-se que cada parâmetro a_i tem distribuição Log-normal com parâmetro $\tau = \left(\mu_\alpha, \sigma_\alpha^2 \right)$. Uma justificativa teórica para a adoção desta distribuição é que na prática os a_i são, em geral, positivos, sugerindo que a distribuição de a_i pode ser modelada por uma distribuição unimodal e com assimetria positiva (Mislevy, 1986), tal como a Log-normal. A transformação $\alpha_i = \log a_i$ resulta em cada α_i tendo uma Distribuição Normal $\left(\mu_\alpha, \sigma_\alpha^2 \right)$, onde

$$\mu_a = \exp \left[\mu_\alpha + \frac{\sigma_\alpha^2}{2} \right] \text{ e } \sigma_a^2 = \left(\exp \left(\sigma_\alpha^2 \right) - 1 \right) \exp \left[2\mu_\alpha + \sigma_\alpha^2 \right].$$

Como a distribuição de a_i é log-normal, sua densidade é :

$$f(a_i | \mu_a, \sigma_a^2) = \frac{1}{\sqrt{2\pi} a_i \sigma_a} \exp \left[-\frac{1}{2\sigma_a^2} (\log a_i - \mu_a)^2 \right].$$

Segue que a segunda parcela de

$$\frac{\partial f^*(\zeta, \eta | u_{..})}{\partial \zeta_i} = \frac{\partial \log L(\zeta, \eta)}{\partial \zeta_i} + \frac{\partial \log f(\zeta)}{\partial \zeta_i} = 0.$$

pode ser escrita como:

$$\frac{\partial \log f(a_i | \mu_a, \sigma_a^2)}{\partial a_i} = -\frac{1}{a_i} \left[1 + \frac{\log a_i - \mu_a}{\sigma_a^2} \right].$$

2.5.5.1.2 Distribuição a priori para b_i

Como os parâmetros de dificuldade estão na mesma escala da proficiência, em geral, supõem-se que cada b_i 's tem Distribuição Normal com vetor de parâmetros $\tau = (\mu_b, \sigma_b^2)$.

Desta forma, a segunda parcela de

$$\frac{\partial f^*(\zeta, \eta | u_{..})}{\partial \zeta_i} = \frac{\partial \log L(\zeta, \eta)}{\partial \zeta_i} + \frac{\partial \log f(\zeta)}{\partial \zeta_i} = 0.$$

pode ser escrita como:

$$\frac{\partial \log f(b_i | \mu_b, \sigma_b^2)}{\partial b_i} = -\frac{(b_i - \mu_b)}{\sigma_b^2}$$

2.5.5.1.3 Distribuição a priori para c_i

Como c_i só pode pertencer ao intervalo $[0, 1]$, uma priori Beta foi proposta por Swaminathan & Gifford (1986).

A função densidade da distribuição Beta com parâmetros $s+1$ e $t+1$ é dada por:

$$f(c_i | s, t) = \frac{\Gamma(s+t+2)}{\Gamma(s+1)\Gamma(t+1)} c_i^s (1-c_i)^t,$$

onde $\Gamma(d)$ é a função Gama, definida por :

$$\Gamma(d) = \int_0^{\infty} x^{d-1} e^{-x} dx.$$

A média desta distribuição é dada por:

$$p = \frac{s+1}{s+t+2},$$

Swaminathan & Gifford propõem, ainda, a seguinte reparametização:

$$\alpha = mp+1 \quad \text{e} \quad \beta = m(1-p)+1,$$

onde:

$$m = s + t + 2.$$

Desta forma, $p = \frac{(s+1)}{m}$, e, conseqüentemente:

$$s = mp-1 \quad \text{e} \quad t = m - s - 2 = m(1-p) - 1.$$

Segue disso que:

$$s = \alpha - 2 \quad \text{e} \quad t = \beta - 2.$$

Retornando-se a :

$$f(c_i|s, t) = \frac{\Gamma(s+t+2)}{\Gamma(s+1)\Gamma(t+1)} c_i^s (1-c_i)^t,$$

obtem:

$$f(c_i|\alpha, \beta) = \frac{\Gamma(\alpha + \beta - 2)}{\Gamma(\alpha - 1)\Gamma(\beta - 1)} c_i^{\alpha-2} (1-c_i)^{\beta-2}.$$

Neste caso, a média ‘p’ passa a ser interpretada como a probabilidade de acerto por indivíduos com baixa proficiência. Desta forma, os parâmetros α e β são definidos para que ‘p’ tenha o valor desejado. Entretanto, Swaminathan & Gifford sugerem que a escolha de ‘m’ deva se situar no intervalo $\{15, \dots, 20\}$, o que leva a uma certa restrição na escolha de α e β .

Para se obter a expressão para a segunda parcela de

$$\frac{\partial f^*(\zeta, \eta|u..)}{\partial \zeta_i} = \frac{\partial \log L(\zeta, \eta)}{\partial \zeta_i} + \frac{\partial \log f(\zeta)}{\partial \zeta_i} = 0,$$

vem:

$$\log f(c_i|\alpha, \beta) = \text{Const} + (\alpha-2)\log c_i + (\beta-2)\log(1-c_i).$$

Conseqüentemente:

$$\frac{\partial \log f(c_i|\alpha, \beta)}{\partial c_i} = \frac{\alpha-2}{c} - \frac{\beta-2}{1-c_i}.$$

Com as componentes

$$\begin{aligned} \frac{\partial \log f(a_i|\mu_a, \sigma_a^2)}{\partial a_i} &= -\frac{1}{a_i} \left[1 + \frac{\log a_i - \mu_a}{\sigma_a^2} \right], \\ \frac{\partial \log f(c_i|\alpha, \beta)}{\partial c_i} &= \frac{\alpha-2}{c} - \frac{\beta-2}{1-c_i}, \end{aligned}$$

tem-se que as equações de estimação para as componentes de ζ_i são:

$$\begin{aligned} a_i &: D(1-c_i) \int_{IR} (\theta - b_i) [r_i(\theta) - P_i f_i(\theta)] w_i d\theta - \frac{1}{a_i} \left[1 + \frac{\log a_i - \mu_a}{\sigma_a^2} \right] = 0, \\ b_i &: -D(1-c_i) \int_{IR} [r_i(\theta) - P_i f_i(\theta)] w_i d\theta - \frac{(b_i - \mu_b)}{\sigma_b^2} = 0, \\ c_i &: \int_{IR} [r_i(\theta) - P_i f_i(\theta)] \frac{w_i}{P_i^*} d\theta + \frac{\alpha-2}{c_i} - \frac{\beta-2}{1-c_i} = 0. \end{aligned}$$

Para efeito de aplicação dos procedimentos iterativos Newton-Raphson ou “Scoring” de Fisher, precisa-se das derivadas segundas destas últimas expressões.

Como as derivadas segundas das primeiras parcelas dessas expressões já foram obtidas na Estimação dos Parâmetros dos Itens, resta apenas a obtenção das segundas parcelas, que são as

seguintes:

$$\begin{aligned}\frac{\partial^2 \log f(a_i | \mu_a, \sigma_a^2)}{\partial a_i^2} &= \frac{1}{a_i \sigma_a^2} [\sigma_a^2 + \log a_i - \mu_a - 1], \\ \frac{\partial^2 \log f(b_i | \mu_b, \sigma_b^2)}{\partial b_i^2} &= \frac{1}{\sigma_b^2}, \\ \frac{\partial^2 \log f(c_i | \alpha, \beta)}{\partial c_i^2} &= -\frac{\alpha - 2}{c_i^2} - \frac{\beta - 2}{(1 - c_i)^2}.\end{aligned}$$

2.5.5.1.4 Equações de estimação em Forma de Quadratura

Sendo conhecidos os nós $\bar{\theta}_k$ e os pesos A_k , $k = 1, 2, \dots, q$, tem-se que as equações de estimação em forma de quadratura para os parâmetros a_i , b_i e c_i são, respectivamente:

$$\begin{aligned}a_i &: D(1 - c_i) \sum_{k=1}^q (\bar{\theta}_k - b_i) [r_{ki} - P_{ki} f_{ki}] W_{ki} - \frac{1}{a_i} \left[1 + \frac{\log a_i - \mu_a}{\sigma_a^2} \right] = 0, \\ b_i &: -D a_i (1 - c_i) \sum_{k=1}^q [r_{ki} - P_{ki} f_{ki}] W_{ki} - \frac{(b_i - \mu_b)}{\sigma_b^2} = 0, \\ c_i &: \sum_{k=1}^q [r_{ki} - P_{ki} f_{ki}] \frac{W_{ki}}{P_{ki}^*} + \frac{\alpha - 2}{c_i} - \frac{\beta - 2}{1 - c_i} = 0.\end{aligned}$$

2.5.5.2 Estimação Bayesiana das habilidades/proficiências

Tal como na estimação por máxima verossimilhança marginal, a estimação bayesiana das proficiências é feita em uma segunda etapa, considerando os parâmetros dos itens fixos. Através da suposição de independência entre as proficiências de diferentes indivíduos, podem-se fazer as estimações em separado para cada indivíduo.

Assumindo que a distribuição a priori para θ_j , $j = 1, 2, \dots, n$, é Normal com vetor de parâmetros $\eta = (\mu, \sigma^2)$ conhecidos. A posteriori para a proficiência do indivíduo j pode ser escrita como:

$$g_j^*(\theta_j) \equiv g(\theta_j | u_j, \zeta, \eta) \propto P(u_j | \theta_j, \zeta) g(\theta_j | \eta).$$

Novamente, pode-se adotar alguma característica de $g_j^*(\theta_j)$ como estimador de θ_j , sendo que as mais adotadas são a média e a moda. A seguir, trata-se da obtenção de cada uma destas

características.

2.5.5.2.1 Estimação das Proficiências pela Moda da Posteriori – MAP

A estimação pela moda da posteriori (ou MAP: *maximum a posteriori*) consiste em obter o máximo de $g_j^*(\theta_j) \equiv g(\theta_j | u_{j.}, \zeta, \eta) \propto P(u_{j.} | \theta_j, \zeta) g(\theta_j | \eta)$.

Por facilidade, trabalha-se com o logaritmo da posteriori,

$$\log g_j^*(\theta_j) = \text{Const} + \log P(u_{j.} | \theta_j, \zeta) + \log g(\theta_j | \eta)$$

Segue que a equação de estimação para θ_j é:

$$\frac{\partial \log g_j^*(\theta_j)}{\partial \theta_j} = \frac{\partial \log P(u_{j.} | \theta_j, \zeta)}{\partial \theta_j} + \frac{\partial \log g(\theta_j | \eta)}{\partial \theta_j} = 0.$$

Pela independência local, tem-se que:

$$\log P(u_{j.} | \theta_j, \zeta) = \log \left[\prod_{i=1}^I P(u_{ji} | \zeta_i, \theta_j) \right] = \sum_{i=1}^I \log P(u_{ji} | \zeta_i, \theta_j).$$

Portanto:

$$\frac{\partial \log P(u_{j.} | \theta_j, \zeta)}{\partial \theta_j} = \sum_{i=1}^I \frac{\partial \log P(u_{ji} | \zeta_i, \theta_j)}{\partial \theta_j} = \sum_{i=1}^I \frac{\partial P(u_{ji} | \zeta_i, \theta_j)}{P(u_{ji} | \zeta_i, \theta_j)} \frac{\partial \theta_j}{\partial \theta_j}.$$

Adotando a priori Normal $\left(\mu, \sigma^2 \right)$ para θ_j , a segunda parcela de

$$\frac{\partial \log g_j^*(\theta_j)}{\partial \theta_j} = \frac{\partial \log P(u_{j.} | \theta_j, \zeta)}{\partial \theta_j} + \frac{\partial \log g(\theta_j | \eta)}{\partial \theta_j} = 0,$$

é :

$$\frac{\partial \log g(\theta_j | \eta)}{\partial \theta_j} = -\frac{(\theta_j - \mu)}{\sigma^2}.$$

Por

$$\begin{aligned} \frac{\partial \log P(u_{j.} | \theta_j, \zeta)}{\partial \theta_j} &= D \sum_{i=1}^I a_i (1 - c_i) (u_{ji} - P_{ji}) w_{ji} \text{ e} \\ \frac{\partial \log g(\theta_j | \eta)}{\partial \theta_j} &= -\frac{(\theta_j - \mu)}{\sigma^2}, \end{aligned}$$

tem-se que a equação de estimação para θ_j é:

$$\theta_j : D \sum_{i=1}^I a_i (1 - c_i) (u_{ji} - P_{ji}) W_{ji} - \frac{(\theta_j - \mu)}{\sigma^2} = 0.$$

Como esta equação não tem solução explícita, pode-se aplicar algum método iterativo para resolvê-la. Para isso será necessária a derivada segunda de $\log g(\theta_j | u_{j.}, \zeta, \eta)$ com relação a θ_j , cuja expressão é:

$$H(\theta_j) = \sum_{i=1}^I (u_{ji} - P_{ji}) W_{ji} \{H_{ji} - (u_{ji} - P_{ji}) W_{ji} h_{ji}^2\} - \frac{1}{\sigma^2}.$$

Onde h_{ji} e H_{ji} são dados por :

$$h_{ji} = (P_{ji}^* Q_{ji}^*)^{-1} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right) = D a_i (1 - c_i), \text{ e,}$$

$$H_{ji} = (P_{ji}^* Q_{ji}^*)^{-1} \left(\frac{\partial^2 P_{ji}}{\partial \theta_j^2} \right) = D^2 a_i^2 (1 - c_i) (1 - 2 P_{ji}^*),$$

respectivamente. Para utilizar-se do método “*Scoring*” de Fisher, deve-se tomar a esperança da expressão acima, resultando em:

$$\Delta(\theta_j) = - \sum_{i=1}^I P_{ji}^* Q_{ji}^* W_{ji} h_{ji}^2 - \frac{1}{\sigma^2}.$$

2.5.5.2.2 Estimação da Proficiência pela Média da Posteriori – EAP

A estimação de θ_j pela média da posteriori (ou EAP: *expected a posteriori*) consiste em obter a esperança da posteriori, que pode ser escrita como:

$$g(\theta | u_{j.}, \zeta, \eta) = \frac{P(u_{j.} | \theta, \zeta) g(\theta | \eta)}{P(u_{j.} | \zeta, \eta)}.$$

Segue que a esperança da posteriori é :

$$\hat{\theta}_j \equiv E[\theta | u_{j.}, \zeta, \eta] = \frac{\int_{IR} \theta g(\theta | \eta) P(u_{j.} | \theta, \zeta) d\theta}{\int_{IR} g(\theta | \eta) P(u_{j.} | \theta, \zeta) d\theta}.$$

Esta forma de estimação para a proficiência tem a vantagem de ser calculada diretamente, não necessitando da aplicação de métodos iterativos. Além disso, as quantidades necessárias para o seu cálculo são um produto final da etapa de estimação. Por conta disso alguns autores recomendam esta escolha para a estimação das proficiências, (Andrade 2000, p.67-75).

3. MATERIAL E MÉTODO

3.1 Método

3.1.1 Introdução ao Software BILOG

Na aplicação da Teoria da Resposta ao Item existe a dificuldade de cálculo da estimação dos parâmetros do modelo. Evidentemente, isto deve ser feito por meio de um programa computacional. Isto ocorre devido não só à grande quantidade de dados que exigem compilação, como também pela complexidade das operações. Existem vários programas que executam os procedimentos da TRI, sendo o mais popular o software conhecido como BILOG, produzido pela *Scientific Software, Inc*, Mislevy, R. J. e Bock, R. D, 1990. Outro programa muito utilizado é o BICAL, desenvolvido no Laboratório de Estatística da Universidade de Chicago.

Na fase da calibração dos itens, como já visto a estimação dos parâmetros é feita pelo Método da Máxima Verossimilhança (MML), admitindo-se a independência entre as respostas dos examinandos.

O programa BILOG que foi utilizado neste trabalho, permite analisar respondentes provenientes de uma única população, e, tem como entrada um arquivo em linguagem própria, extensão.blg, como por exemplo: PROVÃO2003.BLG. Este arquivo/programa informa ao programa principal BILOG:

a) As especificações de entrada de dados formatados em arquivo texto, ou seja, DFName='C:\UNICENTRO\PROVÃOMATEMÁTICA2003.prn'.

b) Salvando os resultados de saída:

SAVE; são as especificações de salvamento dos arquivos de saída, ou seja:

Um arquivo com as estimativas dos parâmetros (PROVÃO2003.par);

Um arquivo com os escores (número absoluto de acertos e percentual) e proficiência de todos os examinandos que compõem o arquivo de dados (PROVÃO2003.sco);

Um arquivo com os gráficos da Curva Característica do Item para cada um dos examinandos (PROVÃO2003.plt).

Um arquivo mostrando os acertos e erros dos examinandos por item PROVÃO2003.sor).
 PARm='C:\ UNICENTRO\PROVÃOMATEMÁTICA2003\ PROVÃO2003.par',
 SCORe='C:\ UNICENTRO\PROVÃOMATEMÁTICA2003\ PROVÃO2003.sco',
 GRAPh='C:\ UNICENTRO\PROVÃOMATEMÁTICA2003\ PROVÃO2003.plt',
 SORTed='C:\ UNICENTRO\PROVÃOMATEMÁTICA2003\ PROVÃO2003.sor';

b) O programa fornece como resultados de saída:

- Os gráficos com as Curvas Características dos Itens em (PROVÃO2003.plt)
- Uma tabela com as estatísticas clássicas da avaliação de itens em (PROVÃO2003.ph1);
- Uma tabela com o número absoluto de acertos e de erros do examinando, percentual estimativa da proficiência e o seu desvio padrão correspondente (PROVÃO2003.ph3 e PROVÃO2003.sco);
- Uma tabela com as estimativas dos parâmetros dos modelos probabilísticos da Teoria da Resposta do Item (PROVÃO2003.par);
- Relatórios com todas as fases do processamento: fase 1, fase 2 e fase 3.

3.1.1.1 Descrição das fases das análises dos dados

O programa executa a análise em três etapas, chamadas de fases 1, 2, e 3, que se caracterizam pelo tipo de tarefas realizadas em cada uma delas.

Na **fase 1**, que é a fase de entrada e leitura de dados, deve-se fornecer ao programa basicamente dois tipos de informação: a identificação de cada indivíduo com suas respectivas respostas já corrigidas, ou seja, já codificadas com 0 ou 1. Nesse caso não há a necessidade do gabarito, pois o programa irá interpretar 1 como acerto e 0 como erro. Essas informações devem estar em arquivos do tipo ASCII. Os arquivos de saída, fornecidos ao usuário, também estarão neste formato. São calculadas algumas estatísticas descritivas, tais como: número de indivíduos submetidos a cada item, número e porcentagem de acerto em cada item e algumas correlações de interesses, como as correlações bisserial e ponto bisserial.

A importância dessa etapa do processamento, além da verificação de que a leitura dos dados foi feita corretamente, é que estas estatísticas são utilizadas posteriormente como valores iniciais para os processo de estimação realizado nas fases seguintes. Além disso, estatísticas

como a correlação bisserial, fornecem um diagnóstico preliminar dos itens, servindo por exemplo, na identificação de itens com problemas no gabarito.

A **fase 2** é a fase da calibração dos itens. Nesta fase, são estimados os parâmetros dos itens, com seus respectivos erros padrão. O BILOG fornece ainda gráficos contendo algumas informações de interesse, tais como as curvas características e as curvas de informação de cada item e do teste. Junto com a curva característica de cada item é fornecido também um teste de ajuste do modelo utilizado.

A **fase 3** é a fase da estimação das proficiências dos respondentes. Aqui são estimadas as proficiências de cada um dos indivíduos, a partir dos resultados obtidos na fase anterior. Essas proficiências inicialmente são estimadas na escala dos parâmetros dos itens. No entanto, pode-se especificar alguns tipos de mudanças na escala, que serão feitas tanto nas proficiências como nos parâmetros estimados na fase anterior.

3.1.2 Métodos para a Calibração dos Itens

O programa realiza inicialmente a calibração (estimação dos parâmetros) dos itens e depois a estimação das proficiências dos respondentes. Dois métodos de estimação para os parâmetros dos itens estão implementados: a máxima verossimilhança marginal e um método bayesiano de estimação por maximização da distribuição marginal a posteriori. Para que os parâmetros dos itens possam ser estimados através de qualquer um desses dois métodos, é necessária a utilização de distribuições de probabilidade para as proficiências dos respondentes. Esses programas assumem que os respondentes representam uma amostra aleatória de uma população de proficiências que pode ser assumida como tendo ou uma distribuição normal padrão, ou ainda uma distribuição discreta arbitrariamente especificada pelo usuário, ou ainda uma distribuição empírica, a ser estimada conjuntamente com os parâmetros dos itens. Esta distribuição empírica é representada na forma de uma distribuição discreta, através de pontos de quadratura.

Na estimação por maximização da distribuição marginal a posteriori, distribuições a priori são definidas para os parâmetros dos itens, o usuário pode especificar prioris normais para o parâmetro de dificuldade, prioris log-normais para os parâmetros de discriminação e prioris beta para o parâmetro de acerto casual.

O BILOG utiliza duas formas de resolver as equações de verossimilhança marginal o algoritmo EM e o método “Scoring” de Fisher, descritos matematicamente neste trabalho.

3.1.3 Métodos Implementados para a Estimação das Proficiências

Uma vez terminada a fase da calibração dos parâmetros, será feita a estimação das habilidades/proficiências dos respondentes. O BILOG tem implementado os 3 métodos de estimação das proficiências (habilidades), isto quando os parâmetros dos itens forem conhecidos, são eles: por máxima verossimilhança, por esperança a posteriori (EAP) e por máximo a posteriori (MAP). No método da máxima verossimilhança, as estimativas das proficiências dos respondentes são calculadas pelo método de Newton-Raphson, utilizando-se uma transformação linear do logito do percentual de acertos dos indivíduos como valores iniciais. Os problemas já descritos com as estimativas dos respondentes que tiveram erro total ou acerto total são contornados através de um artifício: os alunos que erraram todos os itens ganham um meio certo no item mais fácil. Alunos que acertaram todos os itens, perdem um meio certo no item mais difícil. Apesar dessas alternativas implementadas pelos dois programas, este método nem sempre fornece boas estimativas neste caso.

No método EAP, as estimativas para as proficiências são calculadas utilizando-se pontos de quadratura para aproximar a distribuição a priori das proficiências de cada respondente. O número de pontos de quadratura é definido pelo usuário, que pode também escolher entre uma priori que seja normal (e cujos parâmetros podem ser especificados pelo usuário), ou uma distribuição discreta arbitrária (fornecida pelo usuário), ou ainda uma distribuição discreta empírica, através do uso dos pontos de quadratura e de seus respectivos pesos gerados na fase 2.

As estimativas EAP para as proficiências dos respondentes estão sempre definidas, qualquer que seja o padrão de respostas. Além disso, quando se utiliza a estimação por EAP, é fornecida uma estimativa da distribuição de proficiência da população de respondentes, na forma de uma distribuição discreta, dada pelos pontos de quadratura. Esta distribuição é obtida acumulando-se as densidades a posteriori de todos os sujeitos em cada ponto de quadratura. As somas são então normalizadas para obterem-se as probabilidades estimadas em cada ponto. Também são fornecidos a média e o desvio-padrão para essa distribuição estimada.

No método MAP, as estimativas das proficiências são calculadas pelo método Newton-Gauss. Este procedimento sempre converge e fornece estimativas para todos os padrões de resposta possíveis. É assumida uma distribuição a priori normal, cujos parâmetros podem ser especificados pelo usuário, sendo que o padrão definido nesses programas é a normal padrão, (Valle, 1999, págs. 70-72).

3.2 Materiais

A coleta de dados para a presente dissertação foi obtida através de consultas aos arquivos do MEC/INEP para o Curso de Matemática da Universidade Estadual do Centro-Oeste-UNICENTRO, Guarapuava-Paraná. Tanto o relatório do Exame Nacional de Cursos, bem como os gabaritos dos ENC/Provões de Matemática do período de 2000 a 2003, foram coletados na página eletrônica do INEP, conforme os anexo no final deste trabalho.

O ENC/Provão analisado é constituído por um exame de conhecimento de matemática, contendo questões objetivas de múltipla escolha. As questões objetivas destinam-se a avaliar as competências e habilidades desenvolvidas pelos participantes ao longo da escolaridade universitária, a partir de uma Matriz de Competências especialmente desenvolvida para estruturar o ENC/Provão. A tabela abaixo demonstra a cada ano a quantidade de respondentes e de itens que fazem parte da análise.

Tabela 1: Relação da Quantidade de Itens x Respondentes.

Ano do exame:	Quantidade de Itens	Quantidade de Respondentes
2000	25	46
2001	40	59
2002	30	41
2003	40	41

Fonte: O autor

A variável analisada no presente estudo foi à proficiência/habilidade dos examinados, ano a ano e a demonstração é feita através de histogramas, também é determinada as CCIs dos ENC/Provões de 2000 a 2003, com seus respectivos parâmetros métricos, na seção 4 – Resultados e Discussões.

4 RESULTADOS E DISCUSSÕES

4.1 Resultados dos Itens aplicados nos ENC/Provões

As análises estatísticas aplicadas aos dados dos Exames Nacionais de Cursos – ENC/Provões que os alunos da UNICENTRO se submeteram no período de 2000 a 2003, resultou na estimação das proficiências/habilidades (θ) dos respondentes. Usou-se a TRI. Além disso, estimou-se o índice de dificuldade de cada questão e o poder de discriminação das mesmas.

Na análise clássica além dos escores brutos dos testes, obteve-se também a estimação da discriminação dos itens, calculando-se o coeficiente de correlação Bisserial, que é o adequado quando se trabalha com uma variável discreta (a pontuação total no teste) e uma variável dicotômica (indicação de acerto ou erro do item).

4.1.1 Análise do ENC/Provão de Matemática dos alunos da UNICENTRO – 2000

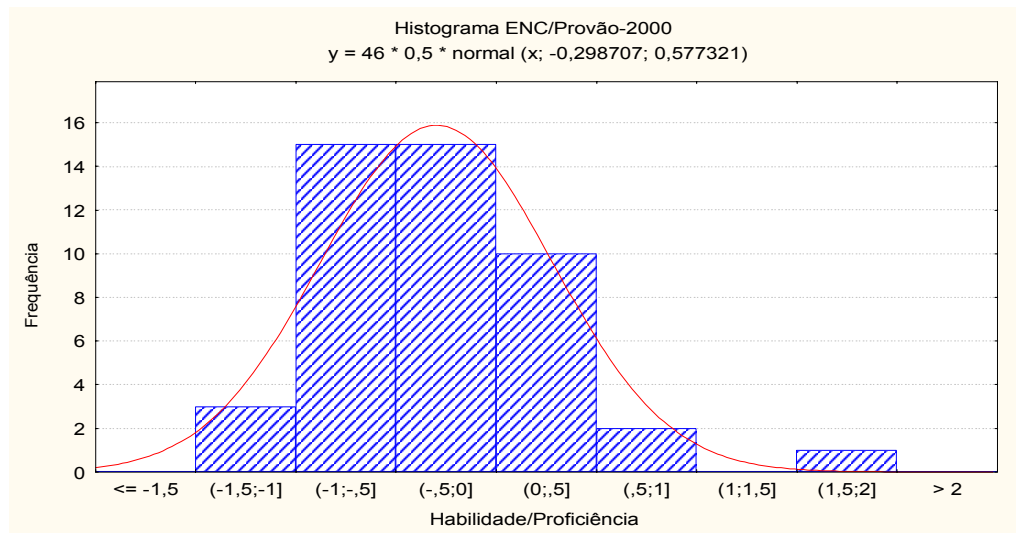
A tabela 2, adiante, mostra a descrição numérica das proficiências/habilidades dos 46 alunos da UNICENTRO, no ano de 2000, examinados no ENC/Provão de matemática. E, a seguir na Figura 5 está o histograma da distribuição de valores. Os valores das proficiências/habilidades estão em escores padronizados, ou seja, estão centrados na média zero e com variância um. Observa-se que a simetria não é boa em torno da média. O modelo Gaussiano ajusta-se muito bem as habilidades, sendo que o valor-p do teste de Gaussianidade Kolmogorov-Smirnov resultou em $p \geq 0,10$. A tabela 3, a seguir mostra os resultados do percentual de acertos e coeficientes de correlação em cada item do teste. Os coeficientes de correlação medem o grau de discriminação de cada item, sendo o mais indicado o Coeficiente de Correlação Bisserial, que para esta avaliação tem um índice máximo de 0,508 e um índice mínimo de $-0,522$. Assim, os testes indicam que os valores estão distribuídos segundo a curva Normal ou de Gauss, o que retrata muito bem a realidade da variabilidade das notas obtidas pelos avaliados.

Tabela 2 – Descrição Numérica das Proficiências/Habilidades

Estatística	Estimativas
Tamanho do Grupo Avaliado	n = 46
Média	$\bar{X} = -0,299$
Variância	$S^2 = 0,333$
Desvio Padrão	$S = 0,577$
Mínimo	$X_{(1)} = -1,239$
Máximo	$X_{(n)} = 1,834$
Amplitude	$R = 3,074$

Fonte: O autor

Figura 5: Histograma das Proficiências/Habilidades, ENC/Provão no ano de 2000.



Fonte: O autor

Tabela 3 – Descrição dos Acertos/Itens e Índices de Discriminação, ENC/Provão 2000

Número do Item	Número Examinandos	Quantidade de acertos	Percentual de acertos	Correlação de Pearson	Correlação Bisserial
01	46	26	0,565	0,262	0,329
02	46	11	0,239	-0,027	-0,038
03	46	13	0,283	0,071	0,094
04	46	14	0,304	0,025	0,033
05	46	20	0,435	0,064	0,081
06	46	08	0,174	0,066	0,097
07	46	02	0,043	0,030	0,065
08	46	10	0,217	0,362	0,508
09	46	12	0,261	0,120	0,163
10	46	09	0,196	0,202	0,290
11	46	12	0,261	0,211	0,285
12	46	02	0,043	-0,237	-0,522
13	46	08	0,174	0,091	0,134
14	46	06	0,130	0,242	0,384
15	46	13	0,283	-0,154	-0,206
16	46	09	0,196	0,152	0,218
17	46	05	0,109	-0,060	-0,101
18	46	07	0,152	-0,025	-0,038
19	46	10	0,217	0,069	0,096
20	46	19	0,413	0,062	0,079
21	46	08	0,174	-0,057	-0,084
22	46	09	0,196	0,127	0,183
23	46	10	0,217	-0,089	-0,125
24	46	14	0,304	0,025	0,033
25	46	06	0,130	-0,014	-0,023

Fonte: O autor

4.1.1.1 Análise da Confiabilidade

A descrição dos escores brutos está na tabela 04 abaixo. Entende-se por escores brutos o número de acertos em um total de 25 itens.

Tabela 04 – Descrição dos Escores Brutos.

Estatística	Valor
Número de Avaliados	n = 46
Média	$\bar{X} = 10,52$
Variância	$S^2 = 29,09333$
Desvio Padrão	S = 5,393824
Mínimo	$X_{(1)} = 1$
Máximo	$X_{(n)} = 14$

Fonte: O autor

A confiabilidade do teste, medida pelo método de Kuder-Richardson, foi de **89,56%**. A expressão de cálculo por este método é a seguinte:

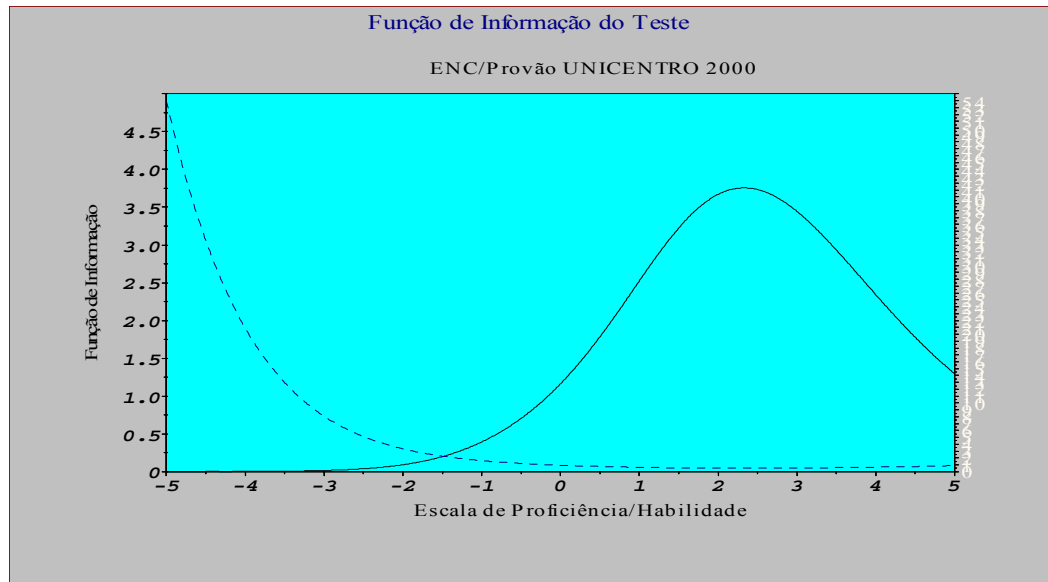
$$\rho_{KR} = \frac{n}{n-1} \left(1 - \frac{\sum_{i=1}^n \theta_i (1 - \theta_i)}{s^2} \right)$$

sendo θ_i o percentual de acerto no item i , s^2 a variância dos escores brutos e n o número de itens.

Os itens que compõem o teste não foram pré-testados, houve somente uma avaliação dos elaboradores das questões quanto ao grau de dificuldade das mesmas (fácil, média, difícil). Assim, o valor observado para a confiabilidade é considerado razoável. O ideal é um índice de confiabilidade acima de 90% (segundo a literatura especializada). O índice obtido está muito próximo do limite mínimo da faixa de valores, considerada ideal, para a confiabilidade e que é de 90% a 100%. Observa-se da tabela anterior que o valor mínimo dos escores brutos foi 01, o máximo de 14 e a média de 10,52 (de um total de 25 itens). Supondo que a avaliação aplicada foi bem elaborada, admite-se que o nível das proficiências/habilidades do grupo dos respondentes é baixa.

A curva da função de informação do teste é apresentada na figura 6.

Figura 6: Função de Informação do Teste – Provão 2000.



Fonte: O autor

Considerando-se os 25 itens aos quais o modelo logístico de três parâmetros foi ajustado e determinado a sua função de informação, pode-se verificar que ela:

- Pode ser adequada se utilizada com examinandos que tenham proficiências/habilidades (θ), a partir de $-1,6$ até $2,2$, isto é, fornece maior quantidade de informações (curva de traço contínuo), nas avaliações de respondentes cuja proficiência/habilidade (θ) esteja neste intervalo.
- Pode ser preciso nas mensurações de proficiência/habilidade (θ) cujos valores estejam no intervalo a partir de $-1,6$ até $2,2$, isto é, para os valores de θ dentro deste valor de intervalo comete-se a menor quantidade de erros de medida (curva de linha tracejada).

As estimativas dos três parâmetros do modelo TRI, a_i , b_i e c_i estão na tabela 05 a seguir, onde o índice de discriminação é representado por 'a', o grau de dificuldade por 'b' e a chance de acerto ao acaso por 'c'.

TABELA 05 - Estimativa dos Parâmetros do Modelo TRI – Provão 2000

Item	a_i	b_i	c_i
01	1.164	-0.161	0.190
02	0.915	2.414	0.180
03	0.870	1.841	0.173
04	0.883	1.660	0.178
05	0.969	0.754	0.203
06	1.026	2.371	0.140
07	0.893	3.696	0.082
08	1.307	1.459	0.135
09	1.063	1.712	0.168
10	1.185	1.849	0.139
11	1.025	1.719	0.165
12	0.854	3.980	0.086
13	0.918	2.520	0.137
14	1.134	2.352	0.115
15	0.757	2.327	0.189
16	1.084	2.117	0.147
17	0.840	3.484	0.119
18	0.840	3.055	0.136
19	0.950	2.409	0.167
20	1.014	0.850	0.201
21	0.954	2.805	0.152
22	1.044	2.132	0.145
23	0.793	2.822	0.169
24	0.795	1.925	0.185
25	0.911	2.988	0.124

Fonte: O autor

Quanto ao grau de dificuldade ‘b’, a melhor amostra segundo a análise e classificação do autor, para os alunos examinados:

- 04 itens excepcionalmente difíceis, são eles os itens: 07, 12, 17 e 18.
- 11 itens difíceis, são eles: 02, 06, 13, 14, 15, 16, 19, 21, 22, 23, e 25, pois o parâmetro do grau de dificuldade de cada um desses itens supera o limite prático superior para a variação do grau de dificuldade que é 2, e os demais itens podem ser considerados:
- 07 itens pouco difíceis, são eles: 03, 04, 08, 09, 10, 11, 24.
- 02 itens fáceis, são eles: 05, 20.
- 01 item muito fácil, é o item: 01.

De modo geral as estimativas médias dos parâmetros dos itens e das proficiências/habilidades dos respondentes do teste são:

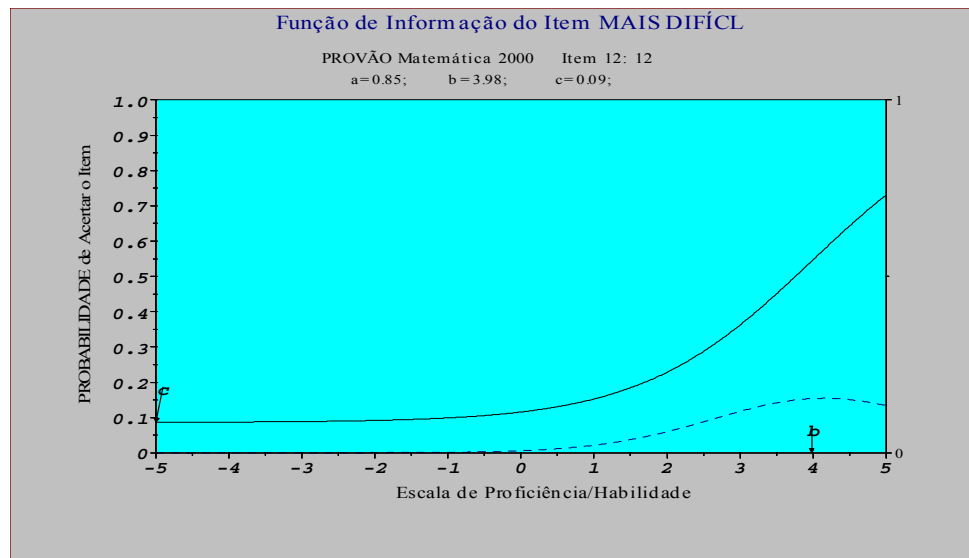
TABELA 06 – Media e Desvio Padrão das Estimativas dos Indicadores – Provão 2000

Indicador	Média	Desvio Padrão
<i>Proficiência/Habilidade</i>	-0,299	0,577
<i>Discriminação 'a'</i>	0,967	0,136
<i>Grau de dificuldade 'b'</i>	2,203	0,925
<i>Acerto ao acaso 'c'</i>	0,153	0,032

Fonte: O autor

Estes são os valores médios verificados, para uma proficiência/habilidade média baixa ' θ ' de -0,299, se tem um valor médio para ' b ' de 2,203 que é muito bom apontando para um teste muito difícil. Já o índice médio de discriminação é 0,967 demonstrando que na média a discriminação é razoável.

Figura 7: Curva Característica do Item 12, (O MAIS DIFÍCIL)



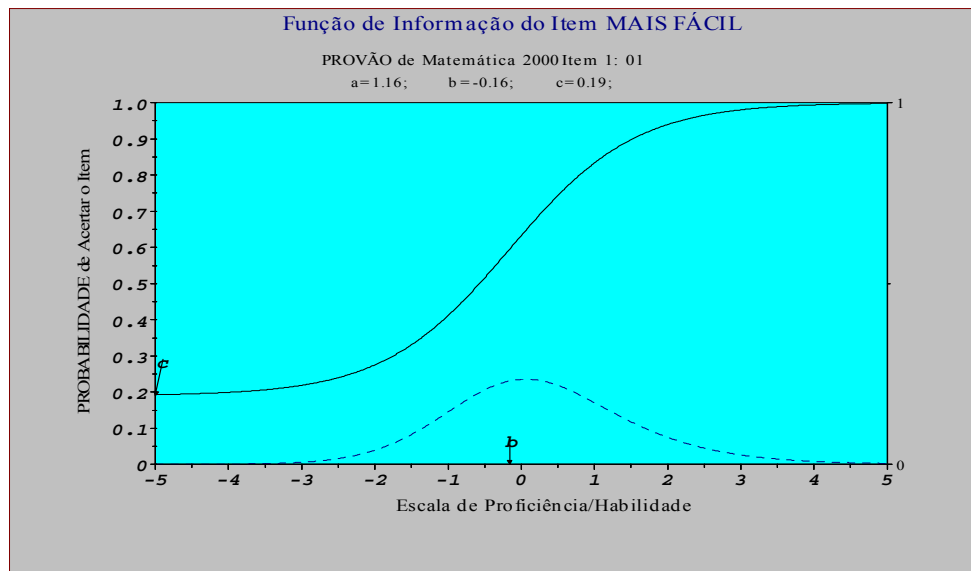
Fonte: O autor

Quanto ao grau de dificuldade, observa-se que o item mais difícil é o 12 com $b_{12} = 3,980$ e o percentual de acerto é 4,30%, pela análise clássica, correspondendo a 2 examinados de um total de 46.

O conteúdo do item mais difícil é de Divisão de Polinômios, ele visa interpretação e a análise de um Polinômio do 4º grau, assim como, a compreensão e utilização de definições, teoremas, propriedades e técnicas matemáticas.

Analisando a forma da CCI, quanto ao índice de discriminação, com valor igual a 0,854, valor próximo da média, observa-se que a curva é muito suave e um deslocamento significativo no eixo da proficiência/habilidade não causa uma alteração significativa na probabilidade de se acertar o item. Isto ocorre devido o item ser muito difícil.

Figura 8: Curva Característica do Item 01, (O MAIS FÁCIL)



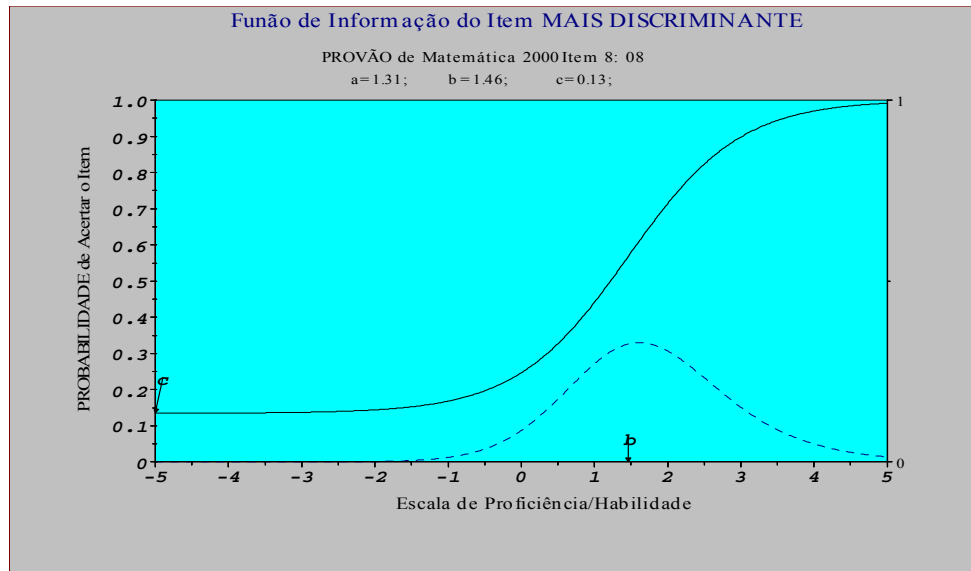
Fonte: O autor

O item mais fácil é o 01, com grau igual a $b_1 = -0,161$ e o percentual de acerto pela análise clássica é de 56,50%, correspondendo a 26 examinados de um total de 46.

O conteúdo do item mais fácil é de Funções Reais Inversas que visa avaliar a capacidade de interpretação gráfica do respondente.

Analisando a CCI quanto ao índice de discriminação com valor igual a 1,164, observa-se que a curva é um pouco íngreme, assim um certo deslocamento no eixo da proficiência/habilidade causa uma alteração significativa na probabilidade de acertar o item.

Figura 9: Curva Característica do Item 08, (O MAIS DISCRIMINANTE)



Fonte: O autor

O parâmetro índice de discriminação ‘a’ dos respondentes, quanto ao conhecimento do assunto varia segundo a tabela 5, de 0,757 a 1,307. Todos os itens têm graus de discriminação de baixos para moderados, significando que discriminam, porém sem exagero. É claro que isto é uma consequência do grau de dificuldade ‘b’ onde a variação “prática” do parâmetro ‘a’ é de 0 a 2 e a do parâmetro ‘b’ é de -2 a +2.

O item mais discriminante é o de número 08 que possui índice igual a 1,307 e na análise clássica tem-se os coeficientes de correlações de Pearson $\rho_8 = 0,362$ e o coeficiente de correlação Bisserial $\rho_{8n} = 0,508$, observa-se que a Curva Característica deste Item é muito íngreme, assim um pequeno deslocamento no eixo da proficiência/habilidade causa uma alteração significativa na probabilidade de acertar o item.

O poder de discriminação deste item é grande, pois é pouco difícil $b_8 = 1,459$ para uma média de dificuldade de 2,203, e além disso, a probabilidade de acerto aumenta significativamente (sobe vertiginosamente) quando se faz um deslocamento na escala das proficiências/habilidades, pode-se notar que a chance de acertá-lo é pequena, $c = 13,50\%$.

O conteúdo do item mais discriminante é Grandezas Proporcionais, o qual visa avaliar a compreensão e a utilização de teoremas e técnicas matemáticas.

4.1.2 Análise do ENC/Provão de Matemática dos alunos da UNICENTRO - 2001

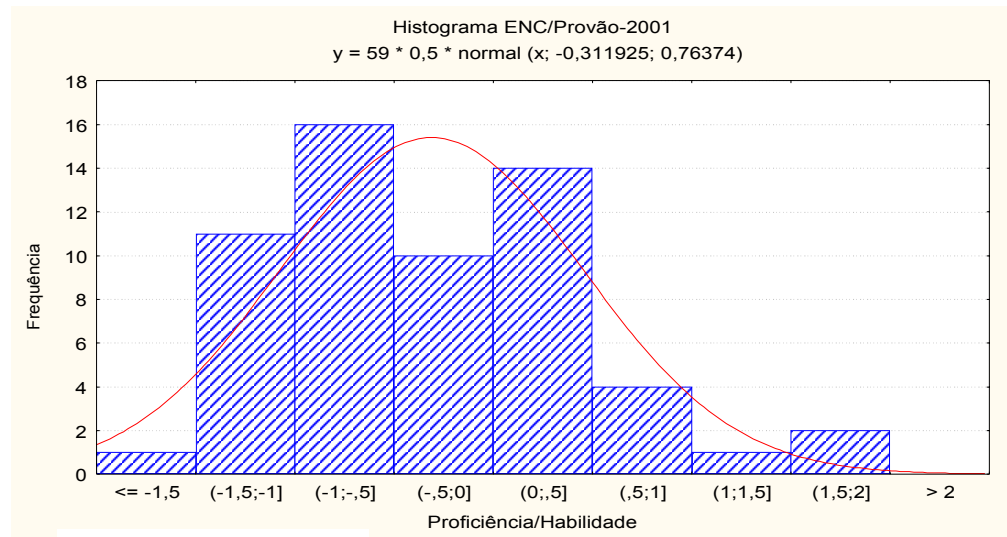
A tabela 7, a seguir, mostra a descrição numérica das proficiências/habilidades dos 59 alunos da UNICENTRO, no ano de 2001, examinados no ENC/Provão de matemática. E, a seguir na Figura 10 está o histograma da distribuição de valores. Os valores das proficiências/habilidades estão em escores padronizados, ou seja, estão centrados na média zero e com variância um. Observa-se que há simetria regular em torno da média. O modelo Gaussiano ajusta-se muito bem as habilidades, sendo que o valor-p do teste de Gaussianidade Kolmogorov-Smirnov resultou em $p < 0,10$. A tabela 8, a seguir mostra os resultados do percentual de acertos e coeficientes de correlação em cada item do teste. Os coeficientes de correlação medem o grau de discriminação de cada item, sendo o mais indicado o Coeficiente de Correlação Bisserial, que para esta avaliação tem um índice máximo de 0,628 e um índice mínimo de $-0,366$. Assim, os testes indicam que os valores estão distribuídos segundo a curva Normal ou de Gauss, o que retrata muito bem a realidade da variabilidade das notas obtidas pelos avaliados

Tabela 7 – Descrição Numérica das Proficiências/Habilidades

Estatística	Estimativas
Tamanho do Grupo Avaliado	$n = 59$
Média	$\bar{X} = -0,312$
Variância	$S^2 = 0,583$
Desvio Padrão	$S = 0,764$
Mínimo	$X_{(1)} = -1,570$
Máximo	$X_{(n)} = 1,930$
Amplitude	$R = 3,500$

Fonte: O autor

Figura 10: Histograma das Proficiências/Habilidades, ENC/Provão no ano de 2001.



Fonte: O autor

Tabela 08 – Descrição dos Acertos/Itens e Índices de Discriminação

Número do Item	Número Examinandos	Quantidade de acertos	Percentual de acertos	Correlação de Pearson	Correlação Bisserial
01	59	29	0,492	0,112	0,140
02	59	16	0,271	0,132	0,177
03	59	18	0,305	0,478	0,628
04	59	11	0,186	0,350	0,509
05	59	23	0,390	0,135	0,172
06	59	10	0,169	0,212	0,315
07	59	27	0,458	0,364	0,457
08	59	38	0,644	0,177	0,228
09	59	31	0,525	0,032	0,040
10	59	14	0,237	0,305	0,420
11	59	14	0,237	0,223	0,307
12	59	08	0,136	-0,037	-0,058
13	59	30	0,508	0,354	0,444
14	59	29	0,492	0,375	0,470
15	59	05	0,085	-0,087	-0,156
16	59	10	0,169	0,133	0,198
17	59	17	0,288	-0,158	-0,210

18	59	25	0,424	0,275	0,346
19	59	37	0,627	0,149	0,190
20	59	18	0,305	0,124	0,162
21	59	04	0,068	-0,191	-0,366
22	59	24	0,407	0,156	0,198
23	59	19	0,322	0,116	0,151
24	59	08	0,136	0,383	0,603
25	59	09	0,153	0,209	0,319
26	59	11	0,186	0,011	0,016
27	59	02	0,034	0,057	0,137
28	59	01	0,017	0,151	0,455
29	59	10	0,169	-0,162	-0,241
30	59	12	0,203	0,064	0,091
31	59	14	0,237	0,346	0,477
32	59	08	0,136	0,011	0,017
33	59	11	0,186	-0,052	-0,075
34	59	16	0,271	0,037	0,050
35	59	26	0,441	0,147	0,185
36	59	33	0,559	0,088	0,111
37	59	20	0,339	0,292	0,377
38	59	05	0,085	-0,102	-0,182
39	59	08	0,136	0,307	0,483
40	59	14	0,237	0,183	0,252

Fonte: O Autor

4.1.2.1 Análise da Confiabilidade

A descrição dos escores brutos está na tabela 09 a seguir. Entende-se por escores brutos o número de acertos em um total de 40 itens.

Tabela 09 – Descrição dos Escores Brutos.

Estatística	Valor
Número de Avaliados	n = 59
Média	$\bar{X} = 16,625$
Variância	$S^2 = 95,72756$
Desvio Padrão	$S = 9,784046$
Mínimo	$X_{(1)} = 4$
Máximo	$X_{(n)} = 24$

Fonte: O Autor.

A confiabilidade do teste, medida pelo método de Kuder-Richardson, foi de **95,04%**. A expressão de cálculo por este método é a seguinte:

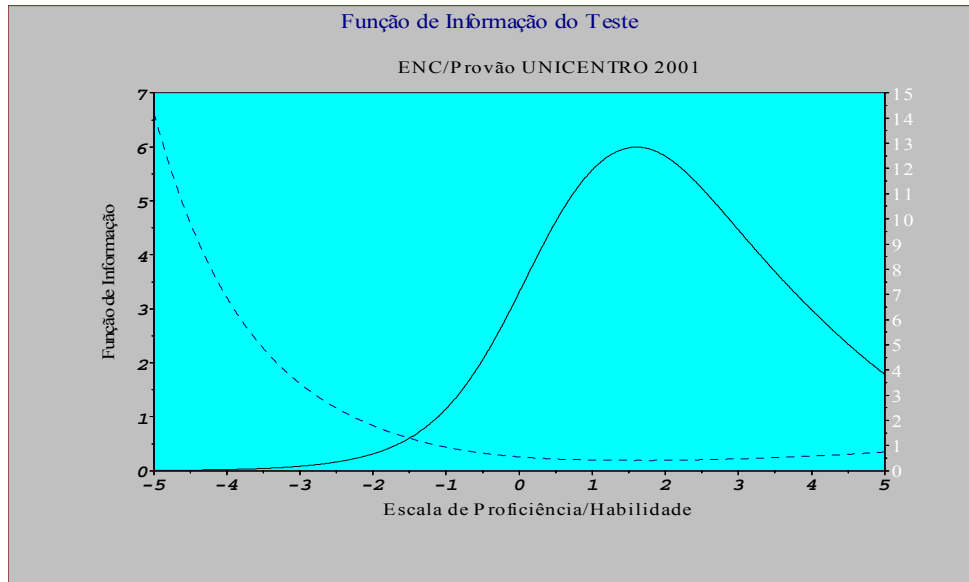
$$\rho_{KR} = \frac{n}{n-1} \left(1 - \frac{\sum_{i=1}^n \theta_i (1 - \theta_i)}{s^2} \right)$$

sendo θ_i o percentual de acerto no item i , s^2 a variância dos escores brutos e n o número de itens.

Os itens que compõem o teste não foram pré-testados, houve somente uma avaliação dos elaboradores das questões quanto ao grau de dificuldade das mesmas (fácil, média, difícil). Assim, o valor observado para a confiabilidade é considerado ótimo. O ideal é um índice de confiabilidade esteja entre 90% a 100% (segundo a literatura especializada). O índice obtido está dentro do limite da faixa de valores. Observa-se da tabela anterior que o valor mínimo dos escores brutos foi 4, o máximo de 24 e a média de 16,62 (de um total de 40 itens). Supondo que a avaliação aplicada foi bem elaborada, admite-se que o nível das proficiências/habilidades do grupo dos respondentes do ENC/Provão de matemática 2001, é baixa.

A curva da função de informação do teste é apresentada na figura 11.

Figura 11: Função de Informação do Teste - Provão 2001.



Fonte: O autor

Considerando-se os 40 itens aos quais o modelo logístico de três parâmetros foi ajustado, e determinado a sua função de informação, pode-se verificar:

- Pode ser adequada se utilizada com examinandos que tenham proficiências/habilidades (θ), a partir de $-1,5$, isto é, fornece maior quantidade de informações (curva de traço contínuo), nas avaliações de respondentes cuja proficiência/habilidade (θ) esteja neste intervalo.
- Pode ser preciso nas mensurações de proficiência/habilidade (θ) cujos valores estejam no intervalo a partir de $-1,5$, isto é, para os valores de θ dentro deste valor de intervalo comete-se a menor quantidade de erros de medida (curva de linha tracejada).

As estimativas dos três parâmetros do modelo TRI, a_i , b_i e c_i estão na tabela 10 a seguir, onde o índice de discriminação é representado por 'a', o grau de dificuldade por 'b' e a chance de acerto ao acaso por 'c'.

TABELA 10 - Estimativas dos Parâmetros de Modelo TRI – Provão 2001

Item	a_i	b_i	c_i
01	0.862	0.446	0.205
02	0.950	1.849	0.170
03	1.821	0.626	0.123
04	1.532	1.492	0.120
05	0.821	1.134	0.189
06	1.195	2.137	0.130
07	1.361	0.250	0.172
08	0.814	-0.557	0.208
09	0.751	0.398	0.223
10	1.127	1.637	0.144
11	0.917	2.052	0.152
12	0.899	3.232	0.127
13	1.267	0.050	0.179
14	1.441	0.058	0.169
15	0.806	3.912	0.100
16	0.944	2.599	0.133
17	0.683	2.855	0.207
18	1.211	0.591	0.191
19	0.826	-0.437	0.208
20	1.079	1.587	0.191
21	0.815	4.137	0.093
22	1.029	0.898	0.203
23	0.904	1.637	0.190
24	1.538	1.695	0.096
25	1.020	2.546	0.124
26	0.894	2.982	0.157
27	0.893	4.021	0.069
28	0.946	3.901	0.057
29	0.738	3.549	0.148
30	0.794	2.691	0.147
31	1.330	1.398	0.139
32	0.819	3.330	0.123
33	0.821	3.103	0.155
34	0.983	1.953	0.181
35	0.888	0.735	0.198
36	0.701	0.171	0.224
37	1.317	0.877	0.164
38	0.842	3.772	0.099
39	1.206	2.139	0.104
40	1.090	1.773	0.151

Fonte: O autor

Quanto ao grau de dificuldade ‘b’, a melhor amostra segundo a análise e classificação do autor, para os alunos examinados:

- 02 itens extraordinariamente difíceis são eles os itens: 21 e 27.
- 07 itens excepcionalmente difíceis são eles os itens: 12, 15, 28, 29, 32, 33 e 38.
- 08 itens difíceis são eles: 06, 11, 16, 17, 25, 26, 30 e 39; pois o parâmetro do grau de dificuldade de cada um desses itens supera o limite prático superior para a variação do grau de dificuldade que é 2, e os demais itens podem ser considerados:
- 10 itens pouco difíceis são eles: 02, 04, 05, 10, 20, 23, 24, 31, 34, 40;
- 11 itens fáceis são eles: 01, 03, 07, 09, 13, 14, 18, 22, 35, 36, 37;
- 02 itens muito fáceis são os itens: 08 e 19.

As estimativas médias dos parâmetros dos itens e das proficiências/habilidades dos respondentes do teste são:

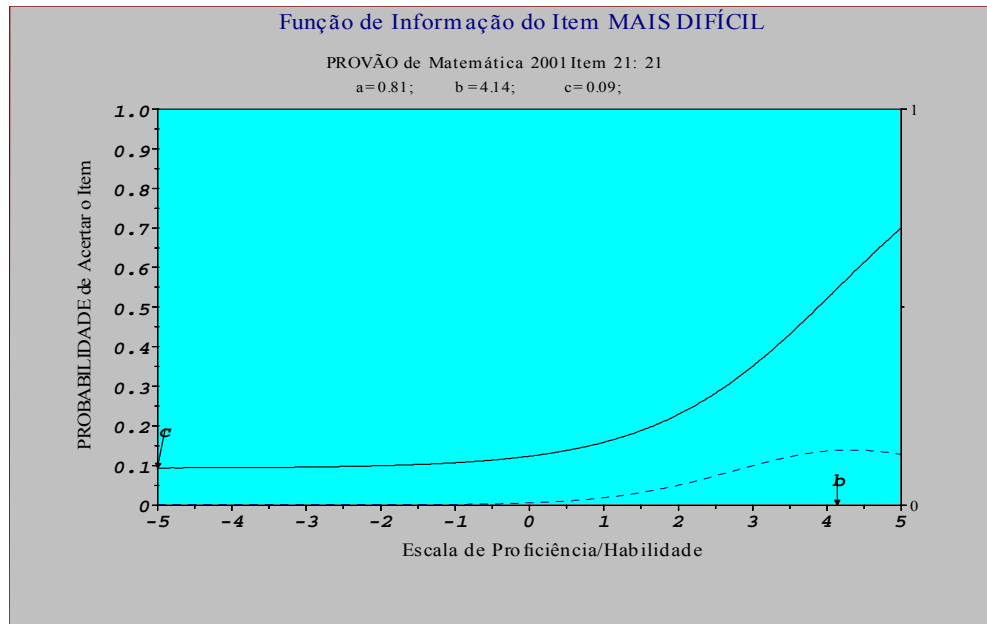
TABELA 11 – Media e Desvio Padrão das Estimativas dos Indicadores – Provão 2001

Indicador	Média	Desvio Padrão
<i>Proficiência/habilidade</i>	-0,312	0,764
<i>Discriminação 'a'</i>	1,022	0,263
<i>Grau de dificuldade 'b'</i>	1,830	1,331
<i>Acerto ao acaso 'c'</i>	0,154	0,043

Fonte: O autor

Estes são os valores médios verificados, para uma proficiência/habilidade média baixa ' θ ' de -0,312, tem-se um valor médio para ' b ' de 1,830 que é muito bom apontando para um teste muito difícil. Já o índice médio de discriminação é 1,022 demonstrando que na média a discriminação do teste é boa.

Figura 12: Curva Característica do Item 21, (O MAIS DIFÍCIL)



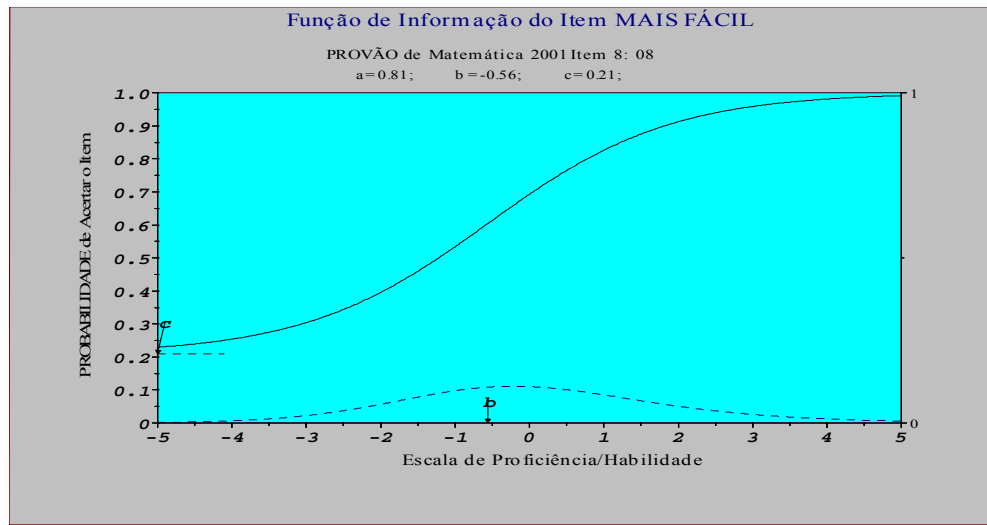
Fonte: O autor

Quanto ao grau de dificuldade, observa-se que o item mais difícil é o 21 com $b_{21} = 4,137$ e o percentual de acerto é 6,80%, pela análise clássica, correspondendo a 4 examinados de um total de 59.

O conteúdo do item mais difícil é de Geometria Espacial, que visa avaliar o raciocínio lógico, a visualização das formas geométricas, a compreensão de definições, teoremas e técnicas matemáticas.

Analisando a forma da CCI, quanto ao índice de discriminação, com valor igual a 0,815, observa-se que a curva é muito suave e um deslocamento significativo no eixo da proficiência/habilidade não causa uma alteração significativa na probabilidade de se acertar o item. Isto ocorre devido o item ser muito difícil.

Figura 13: Curva Característica do Item 08, (O MAIS FÁCIL)



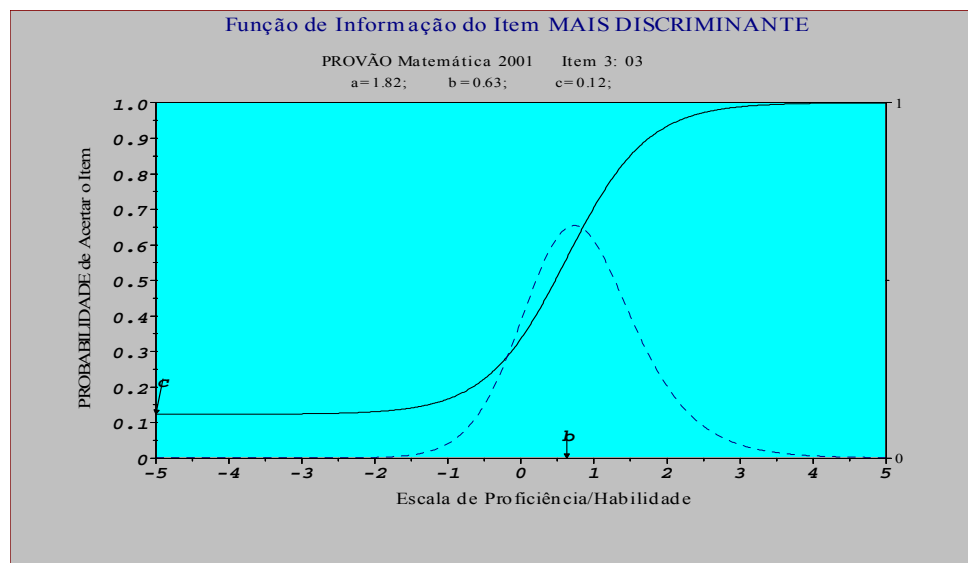
Fonte: O autor

O item mais fácil é o 08, com grau igual a $b_8 = -0,557$ e o percentual de acerto pela análise clássica é de 64,40%, correspondendo a 38 examinados de um total de 59.

O conteúdo do item mais fácil é sobre Números Complexos, o qual permite avaliar a compreensão e a utilização das definições, propriedades e técnicas matemáticas.

Analisando a CCI quanto índice de discriminação com valor igual a 0,814, observa-se que a curva é pouco íngreme, assim um deslocamento no eixo da proficiência/habilidade pode causar uma alteração na probabilidade de acertar o item.

Figura 14: Curva Característica do Item 03, (O MAIS DISCRIMINANTE)



Fonte: O autor

O parâmetro índice de discriminação ‘a’ dos respondentes, quanto ao conhecimento do assunto varia segundo a tabela 10, de 0,683 a 1,821. Todos os itens têm graus de discriminação de baixos para moderados, significando que discriminam, porém sem exagero. É claro que isto é uma consequência do grau de dificuldade.

O item mais discriminante é o de número 03 que possui parâmetro $a_3 = 1,821$ e na análise clássica tem-se os coeficientes de correlações de Pearson é $\rho_3 = 0,478$ e o coeficiente de correlação Bisserial $\rho_{3n} = 0,628$, observa-se que a Curva Característica deste Item é muito íngreme, assim um pequeno deslocamento no eixo da proficiência/habilidade causa uma alteração significativa na probabilidade de acertar o item.

O poder de discriminação muito grande, pois o item é fácil $b_3 = 0,626$ para uma média de dificuldade de 1,830, além disso, a probabilidade de acerto aumenta significativamente (sobe vertiginosamente) quando se faz um deslocamento na escala das proficiências/habilidades, pode-se notar que a chance de acertá-lo é pequena, $c = 12,30\%$. Este é o item mais discriminante nesta análise.

O conteúdo do item mais discriminante é sobre Séries Sequenciais, que visa avaliar a compreensão e a utilização de definições, teoremas, propriedades e técnicas matemáticas.

4.1.3 Análise do ENC/Provão de Matemática dos alunos da UNICENTRO - 2002

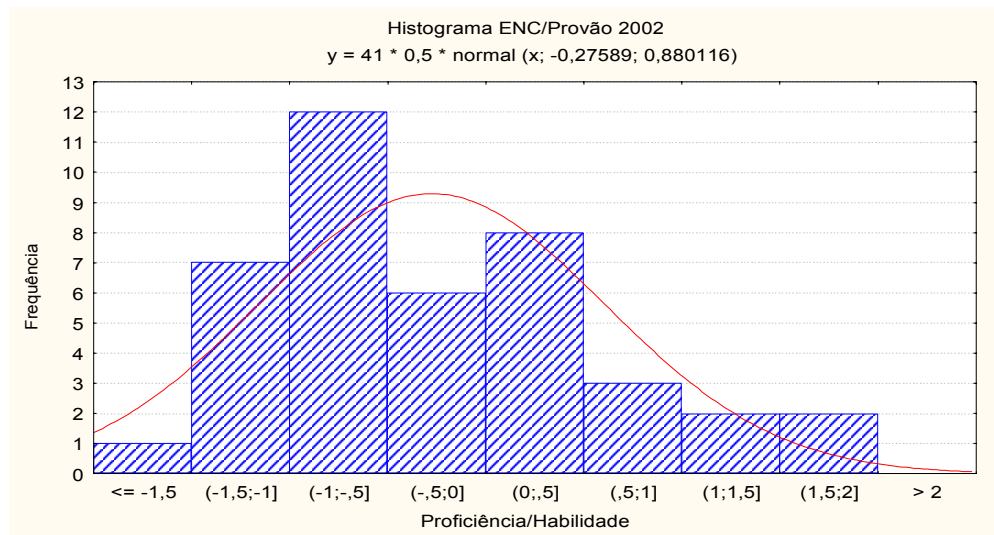
A tabela 12, a seguir, mostra a descrição numérica das proficiências/habilidades dos 41 alunos da UNICENTRO, no ano de 2002, examinados no ENC/Provão de matemática. E, a seguir na Figura 15 está o histograma da distribuição de valores. Os valores das proficiências/habilidades estão em escores padronizados, ou seja, estão centrados na média zero e com variância um. Observa-se uma assimetria regular em torno da média. O modelo Gaussiano ajusta-se muito bem as habilidades, sendo que o valor-p do teste de Gaussianidade Kolmogorov-Smirnov resultou em $p \geq 0,10$. A tabela 13, a seguir mostra os resultados do percentual de acertos e coeficientes de correlação em cada item do teste. Os coeficientes de correlação medem o grau de discriminação de cada item, sendo o mais indicado o Coeficiente de Correlação Bisserial, que para esta avaliação tem um índice máximo de 0,897 e um índice mínimo de $-0,296$. Assim, os testes indicam que os valores estão distribuídos segundo a curva Normal ou de Gauss, o que retrata muito bem a realidade da variabilidade das notas obtidas pelos avaliados.

Tabela 12 – Descrição Numérica das Proficiências/Habilidades

Estatística	Estimativas
Tamanho do Grupo Avaliado	n = 41
Média	$\bar{X} = -0,276$
Variância	$S^2 = 0,775$
Desvio Padrão	S = 0,880
Mínimo	$X_{(1)} = -1,576$
Máximo	$X_{(n)} = 1,952$
Amplitude	R = 3,528

Fonte: O autor

Figura 15: Histograma das Proficiências/Habilidades, ENC/Provão no ano de 2002.



Fonte: O autor

Tabela 13 – Descrição dos Acertos/Itens e Índices de Discriminação

Número do Item	Número Examinandos	Quantidade de acertos	Percentual de acertos	Correlação de Pearson	Correlação Bisserial
01	41	21	0,512	0,331	0,415
02	41	15	0,366	0,484	0,620
03	41	11	0,268	-0,100	-0,135
04	41	17	0,415	0,438	0,553
05	41	23	0,561	0,439	0,553
06	41	21	0,512	0,140	0,175
07	41	15	0,366	0,266	0,341
08	41	07	0,171	0,296	0,436
09	41	10	0,244	0,079	0,108
10	41	03	0,073	0,173	0,324
11	41	26	0,634	0,450	0,576
12	41	12	0,293	-0,212	-0,281
13	41	08	0,195	0,374	0,537
14	41	06	0,146	0,428	0,659
15	41	25	0,610	0,248	0,316
16	41	21	0,512	0,190	0,238
17	41	05	0,122	0,409	0,661
18	41	08	0,195	0,059	0,085
19	41	10	0,244	0,064	0,088
20	41	02	0,049	-0,139	-0,296
21	41	05	0,122	0,142	0,229
22	41	10	0,244	0,281	0,385
23	41	15	0,366	0,253	0,324
24	41	09	0,220	0,113	0,158
25	41	06	0,146	0,373	0,575
26	41	19	0,463	0,295	0,371
27	41	15	0,366	0,442	0,566
28	41	10	0,244	-0,047	-0,065
29	41	02	0,049	0,422	0,897
30	41	36	0,878	0,158	0,255

Fonte: O Autor

4.1.3.1 Análise da Confiabilidade

A descrição dos escores brutos está na tabela 14 abaixo. Entende-se por escores brutos o número de acertos em um total de 30 itens.

Tabela 14 – Descrição dos Escores Brutos.

Estatística	Valor
Número de Avaliados	n = 41
Média	$\bar{X} = 13,10$
Variância	$S^2 = 65,955172$
Desvio Padrão	S = 8,121279
Mínimo	$X_{(1)} = 3$
Máximo	$X_{(n)} = 21$

Fonte: O Autor.

A confiabilidade do teste, medida pelo método de Kuder-Richardson, foi de **95,00%**. A expressão de cálculo por este método é a seguinte:

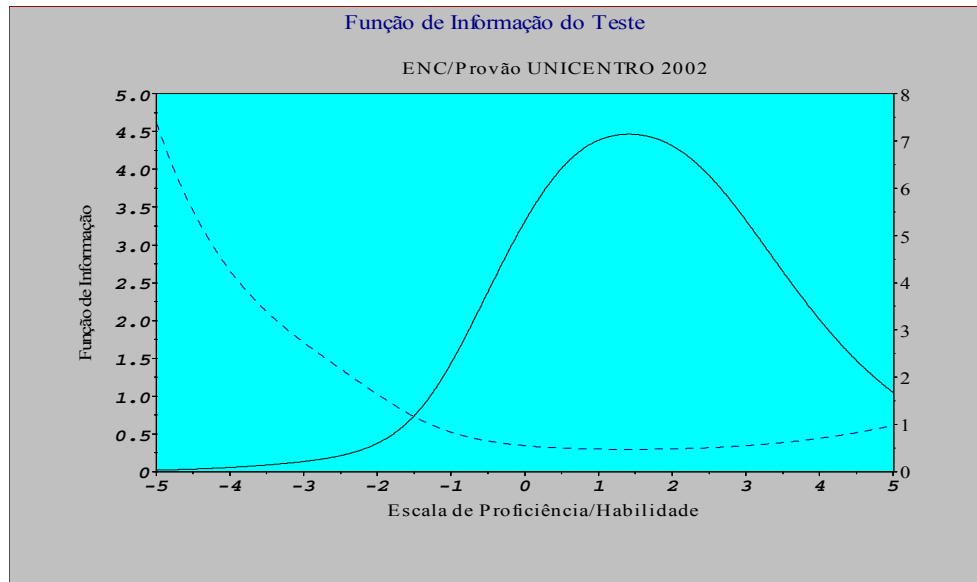
$$\rho_{KR} = \frac{n}{n-1} \left(1 - \frac{\sum_{i=1}^n \theta_i (1 - \theta_i)}{s^2} \right)$$

sendo θ_i o percentual de acerto no item i , s^2 a variância dos escores brutos e n o número de itens.

Os itens que compõem o teste não foram pré-testados, houve somente uma avaliação dos elaboradores das questões quanto ao grau de dificuldade das mesmas (fácil, média, difícil). Assim, o valor observado para a confiabilidade é considerado ótimo. O ideal é um índice de confiabilidade entre 90% a 100% (segundo a literatura especializada). O índice obtido está nesta faixa de valores. Observa-se da tabela anterior que o valor mínimo dos escores brutos foi 3, o máximo de 21 e a média de 13,10 (de um total de 30 itens). Supondo que a avaliação aplicada foi bem elaborada, admite-se que o nível das proficiências/habilidades do grupo dos respondentes do ENC/Provão 2002 é baixa.

A curva da função de informação do teste é apresentada na figura 16.

Figura 16: Função de Informação do Teste - Provão 2002.



Fonte: O autor

Considerando-se os 30 itens aos quais o modelo logístico de três parâmetros foi ajustado e determinado a sua função de informação, pode-se verificar que ela:

- Pode ser adequada se utilizada com examinandos que tenham proficiências/habilidades (θ), a partir de $-1,5$, isto é, fornece maior quantidade de informações (curva de traço contínuo), nas avaliações de respondentes cuja proficiência/habilidade (θ) esteja neste intervalo.
- Pode ser preciso nas mensurações de proficiência/habilidade (θ) cujos valores estejam no intervalo a partir de $-1,5$, isto é, para os valores de θ dentro deste valor de intervalo comete-se a menor quantidade de erros de medida (curva de linha tracejada).

As estimativas dos três parâmetros do modelo TRI, a_i , b_i e c_i estão na tabela 15 a seguir, onde o índice de discriminação é representado por 'a', o grau de dificuldade por 'b' e a chance de acerto ao acaso por 'c'.

TABELA 15 - Estimativas dos Parâmetros de Modelo TRI – Provão 2002

Item	a_i	b_i	c_i
01	1.112	0.147	0.189
02	1.586	0.553	0.149
03	0.695	2.765	0.190
04	1.545	0.372	0.158
05	1.421	-0.202	0.179
06	0.820	0.335	0.201
07	1.430	0.895	0.190
08	1.066	2.155	0.127
09	0.848	2.499	0.178
10	0.890	3.384	0.097
11	1.770	-0.515	0.181
12	0.679	2.797	0.207
13	1.131	1.691	0.118
14	1.340	1.918	0.113
15	1.006	-0.315	0.204
16	0.927	0.347	0.209
17	1.218	2.026	0.095
18	0.897	2.586	0.152
19	0.919	2.304	0.175
20	0.801	4.121	0.094
21	1.003	2.848	0.120
22	1.155	1.719	0.158
23	0.930	1.319	0.197
24	0.797	2.451	0.152
25	1.247	2.036	0.115
26	1.042	0.556	0.202
27	1.328	0.640	0.150
28	0.805	2.785	0.187
29	1.139	2.828	0.077
30	0.827	-2.365	0.212

Fonte: O autor

Quanto ao grau de dificuldade ‘b’, a melhor amostra segundo a análise e classificação do autor, para os alunos examinados:

- 01 item extraordinariamente difícil, ele é o item: 20;
- 01 item excepcionalmente difícil, é o item: 10;
- 12 itens difíceis, são eles: 03, 08, 09, 12, 17, 18, 19, 21, 24, 25, 28, e 29, pois o parâmetro do grau de dificuldade de cada um desses itens supera o limite prático superior para a variação do grau de dificuldade que é 2, e os demais itens podem ser considerados:
- 04 itens pouco difíceis, são eles: 13, 14, 22 e 23;

- 08 itens fáceis, são eles: 01, 02, 04, 06, 07, 16, 26 e 27;
- 04 itens muito fáceis, são eles: 05, 11, 15 e 30.

As estimativas médias dos parâmetros dos itens e das proficiências/habilidades dos respondentes do teste são:

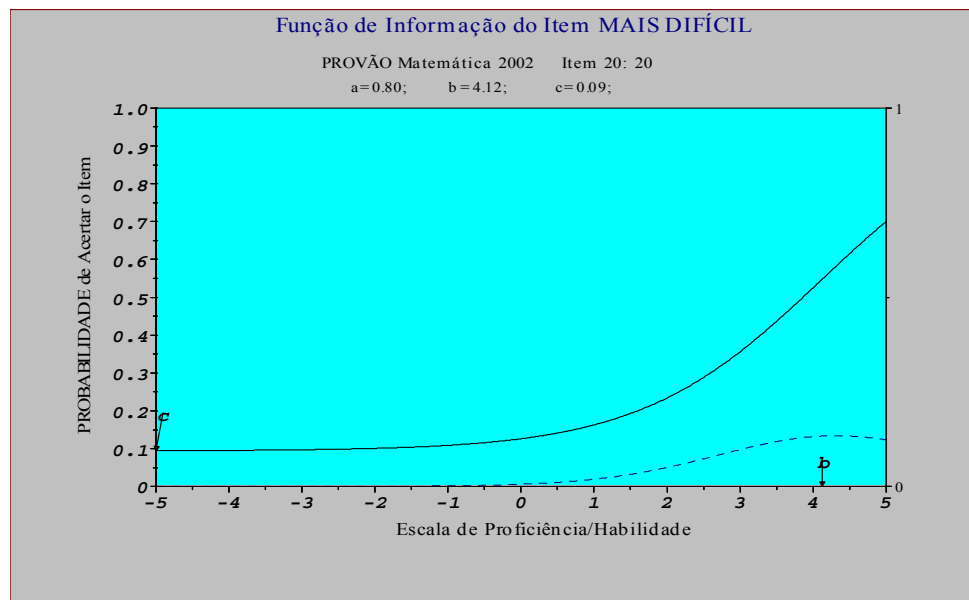
Tabela 16 – Media e Desvio Padrão das Estimativas dos Indicadores – Provão 2002

Indicador	Média	Desvio Padrão
<i>Proficiência/habilidade</i>	-0,276	0,880
<i>Discriminação 'a'</i>	1,079	0,279
<i>Grau de dificuldade 'b'</i>	1,489	1,412
<i>Acerto ao acaso 'c'</i>	0,159	0,040

Fonte: O autor

Estes são os valores médios verificados, para uma proficiência/habilidade média baixa ' θ ' de -0,276, tem-se um valor médio para ' b ' de 1,489 que é muito bom apontando para um teste pouco difícil. Já o índice médio de discriminação é 1,079 demonstrando que na média a discriminação é boa.

Figura 17: Curva Característica do Item 20, (O MAIS DIFÍCIL)



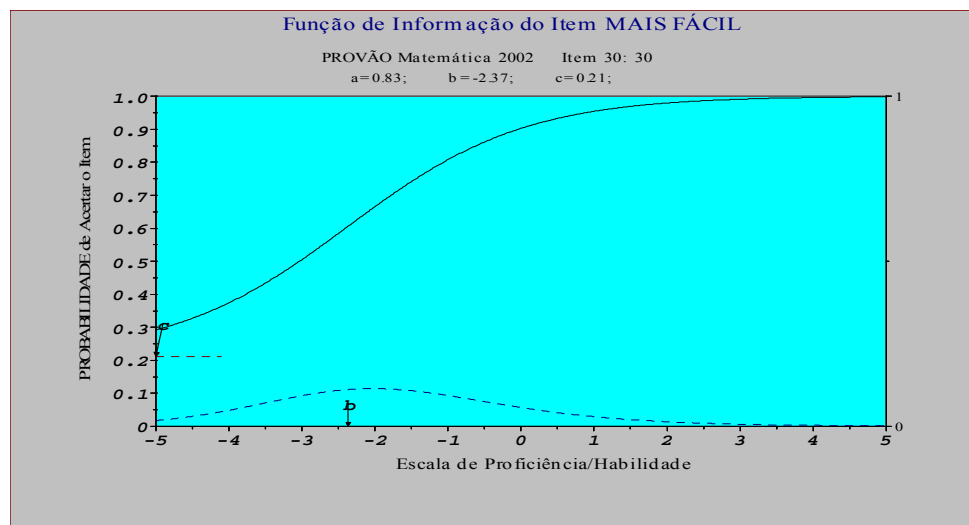
Fonte: O autor

Quanto ao grau de dificuldade, observa-se que o item mais difícil é o 20 com $b_{20} = 4,121$ e o percentual de acerto é 4,90%, pela análise clássica, correspondendo a 2 examinados de um total de 41.

O conteúdo do item mais difícil é Geometria Espacial, que visa avaliar o desenvolvimento do raciocínio lógico, a interpretação e análise de dados, como também a visualização de formas geométricas.

Analisando a forma da CCI, quanto ao índice de discriminação, com valor igual a 0,801, observa-se que a curva é muito suave e um deslocamento significativo no eixo da proficiência/habilidade não causa uma alteração significativa na probabilidade de se acertar o item. Isto ocorre devido o item ser muito difícil.

Figura 18: Curva Característica do Item 30, (O MAIS FÁCIL)



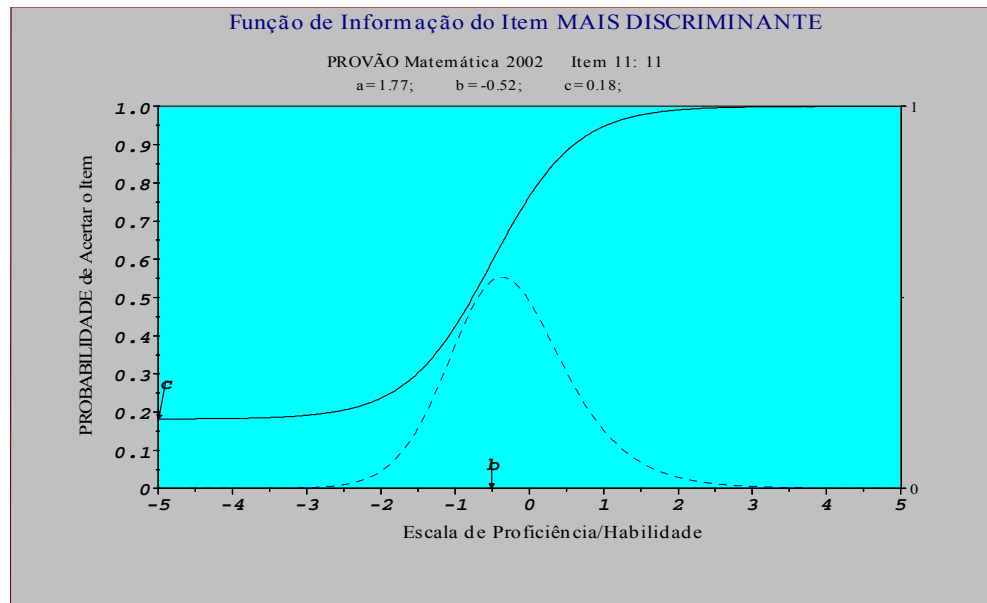
Fonte: O autor

Já o item mais fácil é o 30, com grau igual a $b_{30} = -2,365$ e o percentual de acerto pela análise clássica é de 87,80%, correspondendo a 36 examinados de um total de 41.

O conteúdo do item mais fácil é de Números Racionais, que visa avaliar a compreensão de teoremas, propriedades, conceitos e técnicas matemáticas.

Analisando a CCI quanto ao índice de discriminação com valor igual a 0,827, observa-se que a curva é pouco suave, assim um deslocamento no eixo da proficiência/habilidade não causa uma alteração significativa na probabilidade de acertar o item, logo ele é pouco discriminante em relação aos demais.

Figura 19: Curva Característica do Item 11, (O MAIS DISCRIMINANTE)



Fonte: O autor

O parâmetro índice de discriminação ‘a’ dos respondentes, quanto ao conhecimento do assunto varia segundo a tabela 15, de 0,679 a 1,770. Todos os itens têm graus de discriminação de baixos para moderados, significando que discriminam, porém sem exagero. É claro que isto é uma consequência do grau de dificuldade.

O item mais discriminante é o de número 11 que possui índice igual a 1,770, e na análise clássica tem-se os coeficientes de correlações de Pearson $\rho_{11} = 0,450$ e o coeficiente de correlação Bisserial $\rho_{11n} = 0,576$, observa-se que a Curva Característica deste Item é muito íngreme, assim um pequeno deslocamento no eixo da proficiência/habilidade causa uma alteração significativa na probabilidade de acertar o item.

O poder de discriminação deste item é muito grande, pois o ele é difícil $b_{11} = 2,797$, para uma média de dificuldade de 1,489, e, além disso, a probabilidade de acerto aumenta significativamente (sobe vertiginosamente) quando se faz um deslocamento na escala das proficiências/habilidades, pode-se notar que a chance de acertá-lo é pequena, $c = 18,10\%$. Logo, este é o item mais discriminante nesta análise.

O conteúdo do item mais discriminante é de Critérios de Divisibilidade, que visa avaliar a compreensão e a utilização de definições, teoremas, propriedades e técnicas matemáticas.

4.1.4 Análise do ENC/Provão de Matemática dos alunos da UNICENTRO - 2003

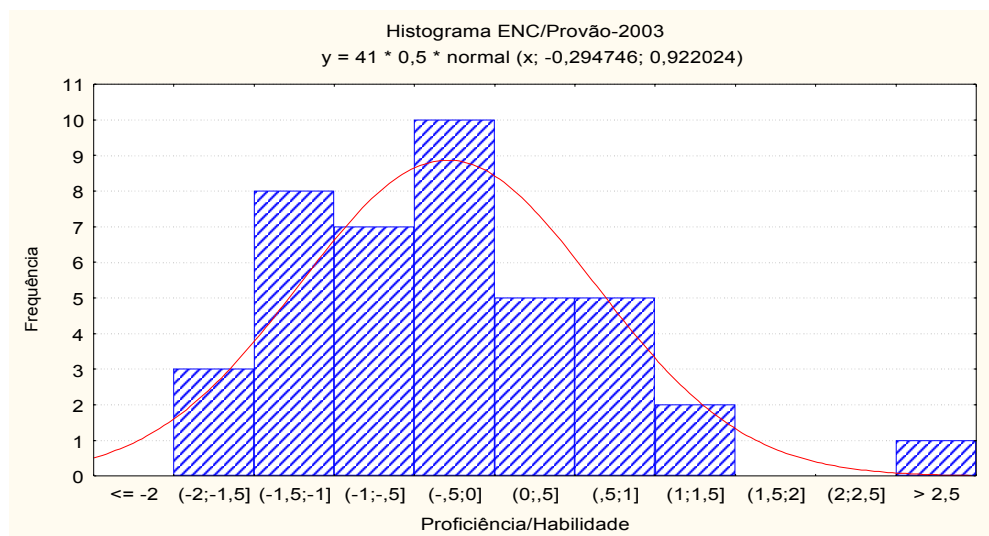
A tabela 17, a seguir, mostra a descrição numérica das proficiências/habilidades dos 41 alunos da UNICENTRO, no ano de 2003, examinados no ENC/Provão de matemática. E, a seguir na Figura 20 está o histograma da distribuição de valores. Os valores das proficiências/habilidades estão em escores padronizados, ou seja, estão centrados na média zero e com variância um. Observa-se que a simetria não é boa em torno da média. O modelo Gaussiano ajusta-se muito bem as habilidades, sendo que o valor-p do teste de Gaussianidade Kolmogorov-Smirnov resultou em $p \geq 0,10$. A tabela 18, a seguir mostra os resultados do percentual de acertos e coeficientes de correlação em cada item do teste. Os coeficientes de correlação medem o grau de discriminação de cada item, sendo o mais indicado o Coeficiente de Correlação Bisserial, que para esta avaliação tem um índice máximo de 1,090 e um índice mínimo de $-0,233$. Assim, os testes indicam que os valores estão distribuídos segundo a curva Normal ou de Gauss, o que retrata muito bem a realidade da variabilidade das notas obtidas pelos avaliados.

Tabela 17 – Descrição Numérica das Proficiências/Habilidades

Estatística	Estimativas
Tamanho do Grupo Avaliado	$n = 41$
Média	$\bar{X} = -0,295$
Variância	$S^2 = 0,850$
Desvio Padrão	$S = 0,922$
Mínimo	$X_{(1)} = -1,963$
Máximo	$X_{(n)} = 2,534$
Amplitude	$R = 4,497$

Fonte: O Autor

Figura 20: Histograma das Proficiências/Habilidades, ENC/Provão no ano de 2003.



Fonte: O autor

Tabela 18 – Descrição dos Acertos/Itens e Índices de Discriminação

Número do Item	Número Examinandos	Quantidade de acertos	Percentual de acertos	Correlação de Pearson	Correlação Bisserial
01	41	13	0,317	0,433	0,565
02	41	35	0,854	0,047	0,072
03	41	36	0,878	0,377	0,609
04	41	13	0,317	0,065	0,085
05	41	19	0,463	0,548	0,688
06	41	05	0,122	0,297	0,480
07	41	23	0,561	0,565	0,711
08	41	16	0,390	0,380	0,483
09	41	01	0,024	0,048	0,128
10	41	15	0,366	0,217	0,278
11	41	22	0,537	-0,017	-0,022
12	41	05	0,122	0,150	0,243
13	41	02	0,049	0,267	0,568
14	41	23	0,561	0,180	0,227
15	41	13	0,317	0,219	0,287
16	41	12	0,293	-0,028	-0,037

17	41	23	0,561	0,441	0,555
18	41	22	0,537	0,222	0,278
19	41	15	0,366	0,434	0,555
20	41	03	0,073	0,513	0,961
21	41	08	0,195	-0,162	-0,233
22	41	27	0,659	0,283	0,365
23	41	12	0,293	0,106	0,141
24	41	09	0,220	0,337	0,471
25	41	21	0,512	0,128	0,161
26	41	31	0,756	-0,088	-0,121
27	41	11	0,268	0,140	0,188
28	41	07	0,171	0,241	0,358
29	41	16	0,390	0,158	0,201
30	41	17	0,415	0,452	0,571
31	41	15	0,366	0,117	0,150
32	41	13	0,317	0,188	0,246
33	41	15	0,366	0,392	0,502
34	41	02	0,049	0,513	1,090
35	41	17	0,415	0,370	0,467
36	41	10	0,244	0,066	0,091
37	41	12	0,293	0,138	0,182
38	41	08	0,195	0,253	0,363
39	41	11	0,268	0,012	0,016
40	41	14	0,341	0,207	0,268

Fonte: O Autor

4.1.4.1 Análise da Confiabilidade

A descrição dos escores brutos está na tabela 19 a seguir. Entende-se por escores brutos o número de acertos em um total de 40 itens

Tabela 19 – Descrição dos Escores Brutos.

Estatística	Valor
Número de Itens	n = 41
Média	$\bar{X} = 14,80$
Variância	$S^2 = 70,984615$
Desvio Padrão	$S = 8,4252368$
Mínimo	$X_{(1)} = 7$
Máximo	$X_{(n)} = 32$

Fonte: O Autor.

A confiabilidade do teste, medida pelo método de Kuder-Richardson, foi de **91,61%**. A expressão de cálculo por este método é a seguinte:

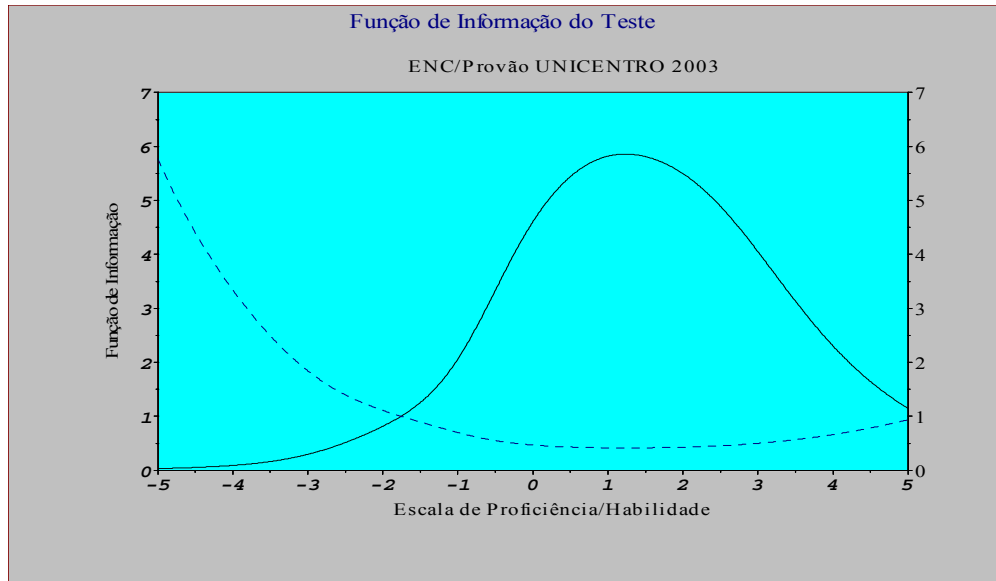
$$\rho_{KR} = \frac{n}{n-1} \left(1 - \frac{\sum_{i=1}^n \theta_i (1 - \theta_i)}{s^2} \right)$$

sendo θ_i o percentual de acerto no item i , s^2 a variância dos escores brutos e n o número de itens.

Os itens que compõem o teste não foram pré-testados, houve somente uma avaliação dos elaboradores das questões quanto ao grau de dificuldade das mesmas (fácil, média, difícil). Assim, o valor observado para a confiabilidade é considerado muito bom. O ideal é um índice de confiabilidade acima de 90% (segundo a literatura especializada). O índice obtido está dentro do limite da faixa de valores, considerada ideal, para a confiabilidade e que é de 90% a 100%. Observa-se da tabela anterior que o valor mínimo dos escores brutos foi 7, o máximo de 32 e a média de 14,80 (de um total de 40 itens). Este valor da confiabilidade indica que nos 40 itens que compõem o teste não existem traços significativos de ambigüidade e textos que conduzam a interpretações diferentes, por partes dos respondentes, quanto à solução. Supondo que a avaliação aplicada foi bem elaborada, admite-se que o nível das proficiências/habilidades do grupo dos respondentes é baixa.

A curva da função de informação do teste é apresentada na figura 21.

Figura 21: Função de Informação do Teste - Provão 2003.



Fonte: O autor

Considerando-se os 40 itens aos quais o modelo logístico de três parâmetros foi ajustado e determinado a sua função de informação, pode-se verificar que ela:

- Pode ser adequada se utilizada com examinandos que tenham proficiências/habilidades (θ), compreendida a partir de $-1,7$, isto é, fornece maior quantidade de informações (curva de traço contínuo), nas avaliações de respondentes cuja proficiência/habilidade (θ) esteja neste intervalo.
- Pode ser preciso nas mensurações de proficiência/habilidade (θ) cujos valores estejam compreendida a partir de $-1,7$, isto é, para os valores de θ compreendido neste intervalo comete-se a menor quantidade de erros de medida (curva de linha tracejada).

As estimativas dos três parâmetros do modelo TRI, a_i , b_i e c_i estão na tabela 20. a seguir, onde o índice de discriminação é representado por 'a', o grau de dificuldade por 'b' e a chance de acerto ao acaso por 'c'.

TABELA 20: Estimativas dos Parâmetros do Modelo TRI – Provão 2003

Item	a_i	b_i	c_i
01	1.509	0.811	0.148
02	0.760	-2.235	0.213
03	1.428	-1.875	0.197
04	0.884	1.991	0.208
05	1.595	0.144	0.166
06	1.142	2.436	0.112
07	2.022	-0.302	0.165
08	1.155	0.684	0.169
09	0.833	3.996	0.076
10	1.163	1.173	0.206
11	0.674	0.310	0.212
12	0.888	2.895	0.114
13	0.878	3.531	0.084
14	0.914	0.089	0.221
15	1.066	1.450	0.186
16	0.844	2.232	0.199
17	1.411	-0.274	0.169
18	1.039	0.047	0.197
19	1.311	0.673	0.158
20	1.261	2.431	0.084
21	0.686	3.421	0.165
22	0.992	-0.652	0.199
23	0.880	2.043	0.192
24	1.126	1.760	0.142
25	0.841	0.360	0.208
26	0.572	-1.508	0.224
27	1.036	1.962	0.184
28	1.047	2.315	0.135
29	0.799	1.228	0.195
30	1.551	0.454	0.177
31	0.865	1.443	0.203
32	1.280	1.323	0.191
33	1.376	0.685	0.163
34	1.357	2.492	0.074
35	1.082	0.575	0.169
36	0.774	2.459	0.168
37	0.893	1.934	0.187
38	1.307	1.794	0.139
39	0.737	2.439	0.180
40	0.943	1.479	0.195

Fonte: O autor

Quanto ao grau de dificuldade ‘b’, a melhor amostra segundo a análise e classificação do autor, para os alunos examinados:

- 03 itens excepcionalmente difíceis, são eles os itens: 09, 13 e 21;
- 09 itens difíceis, são eles: 06, 12, 16, 20, 23, 28, 34, 36 e 39, pois o parâmetro do grau de dificuldade de cada um desses itens supera o limite prático superior para a variação do grau de dificuldade que é 2, e os demais itens podem ser considerados:
- 11 itens pouco difíceis, são eles: 04, 10, 15, 24, 27, 29, 31, 32, 37, 38 e 40;
- 11 itens fáceis, são eles: 01, 05, 08, 11, 14, 18, 19, 25, 30, 33 e 35;
- 06 itens muito fáceis, são eles: 02, 03, 07, 17, 22, e 26.

As estimativas médias dos parâmetros dos itens e das proficiências/habilidades dos respondentes do teste são:

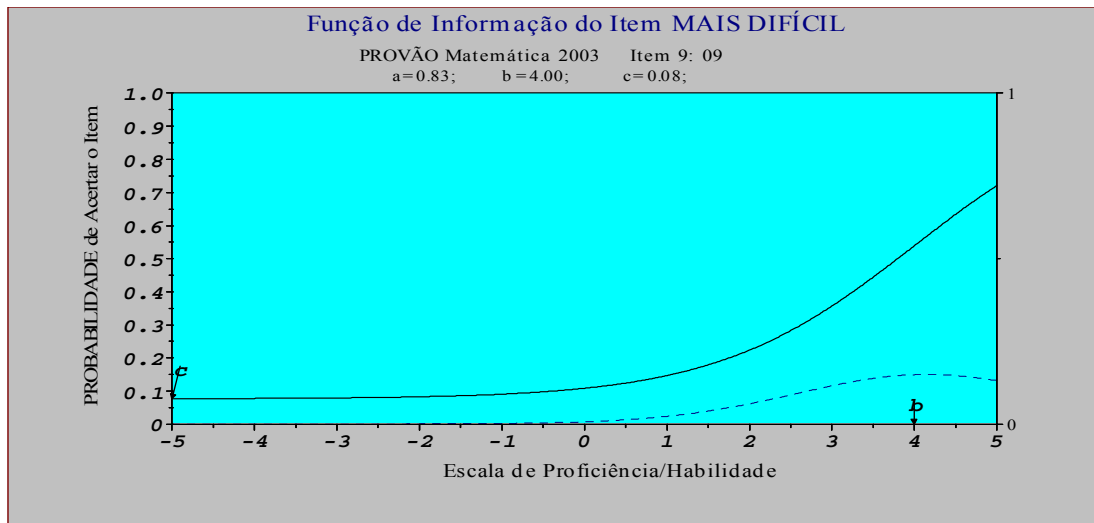
TABELA 21 – Media e Desvio Padrão das Estimativas dos Indicadores – Provão 2003

Indicador	Média	Desvio Padrão
<i>Proficiência/habilidade</i>	-0,295	0,922
<i>Discriminação ‘a’</i>	1,073	0,303
<i>Grau de dificuldade ‘b’</i>	1,205	1,415
<i>Acerto ao acaso ‘c’</i>	0,169	0,040

Fonte: O autor

Estes são os valores médios verificados, para uma proficiência/habilidade média baixa θ de -0,295, tem-se um valor médio para ‘b’ de 1,205 que é bom apontando para um teste pouco difícil. Já o índice médio de discriminação é 1,073 demonstrando que na média a discriminação é boa.

Figura 22: Curva Característica do Item 09, (O MAIS DIFÍCIL)



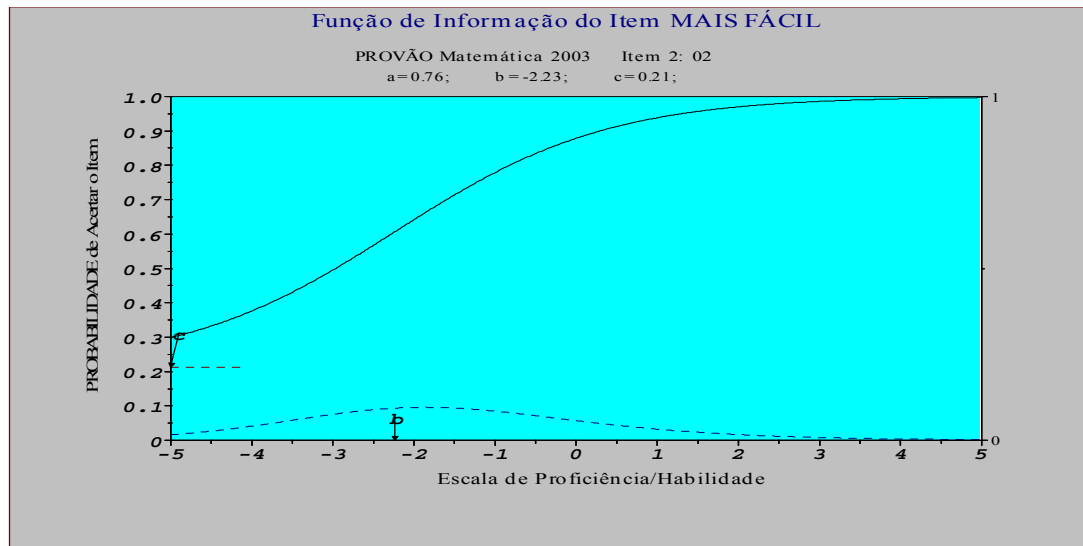
Fonte: O autor

Quanto ao grau de dificuldade, observa-se que o item mais difícil é o 09 com $b_9 = 3,996$ e o percentual de acerto é 2,40%, pela análise clássica, correspondendo a 01 examinado de um total de 41.

O conteúdo do item mais difícil é Álgebra, mais especificamente: Anéis, que visa avaliar a compreensão dos conceitos abstratos e argumentações matemáticas dos respondentes.

Analisando a forma da CCI, quanto ao índice de discriminação, com valor igual a 0,833, observa-se que a curva é muito suave e um deslocamento significativo no eixo da proficiência/habilidade não causa uma alteração significativa na probabilidade de se acertar o item. Isto ocorre devido o item ser muito difícil.

Figura 23: Curva Característica do Item 02, (O MAIS FÁCIL)



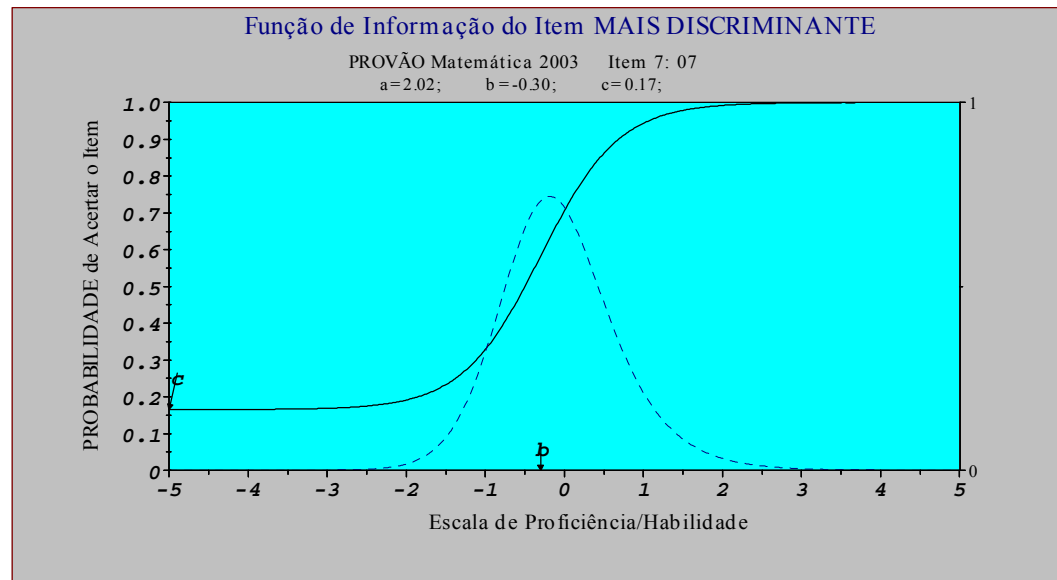
Fonte: O autor

O item mais fácil é o 02, com grau igual a $b_2 = -2,235$ e o percentual de acerto pela análise clássica é de 85,40%, correspondendo a 35 examinados de um total de 41.

O conteúdo do item mais fácil é a busca de soluções reais de uma equação, que visa avaliar nos respondentes a interpretação, análise de conceitos e técnicas matemáticas.

Analisando a CCI quanto ao índice de discriminação com valor igual a 0,760, observa-se que a curva é bastante suave, assim um deslocamento no eixo da proficiência/habilidade não causa uma alteração significativa na probabilidade de acertar o item. Logo este item é pouco discriminante em relação aos demais.

Figura 24: Curva Característica do Item 07, (O MAIS DISCRIMINANTE)



Fonte: O autor

O parâmetro índice de discriminação ‘a’ dos respondentes, quanto ao conhecimento do assunto varia segundo a tabela 20, de 0,572 a 2,022. Todos os itens têm graus de discriminação de baixos para altos, significando que discriminam, porém sem exagero. É claro que isto é uma consequência do grau de dificuldade.

O item mais discriminante é o de número 07 que possui parâmetro $a_7 = 2,022$ e na análise clássica tem-se os coeficientes de correlações de Pearson $\rho_7 = 0,565$ e o coeficiente de correlação Bisserial, é, $\rho_{7n} = 0,711$, novamente ressalta-se que a variação “prática” do parâmetro ‘a’ é de 0 a 2 e a do parâmetro ‘b’ é de -2 a +2, observa-se que a Curva Característica deste Item é muito íngreme, assim um pequeno deslocamento no eixo da proficiência/habilidade causa uma alteração significativa na probabilidade de acertar o item.

O poder de discriminação muito grande, pois o item é muito fácil $b_7 = -0,302$, para uma média de dificuldade de 1,205, e, além disso, a probabilidade de acerto aumenta significativamente quando se faz um deslocamento na escala das proficiências/habilidades. Pode-se notar que, a chance de acertá-lo é pequena, $c = 16,50\%$. Logo, este é o item mais discriminante nesta análise.

O conteúdo do item mais discriminante é de Geometria Plana, que visa avaliar a compreensão e utilização de teoremas, propriedades, conceitos e técnicas matemáticas.

5 CONCLUSÃO

A avaliação em educação de uma maneira geral, segundo a maioria dos pesquisadores, serve para julgar e classificar os respondentes. A participação do avaliado nesse processo é muito pequena e, muitas vezes, nem ele mesmo tem clareza do porquê dos resultados obtidos. A nota chega como sentença, definindo seu destino acadêmico e, às vezes, até seu destino fora da IES. E, muitas vezes revela-se uma total ausência de reflexão sobre o significado da avaliação.

A construção de um banco de questões pré-testadas e com informações quanto ao grau de dificuldade, grau de discriminação e outras informações poderiam ser extremamente úteis à medida da qualidade da aprendizagem de uma disciplina. Isto, sem dúvida, deverá fazer parte do moderno planejamento e controle educacional.

Este trabalho procurou apresentar um estudo do método da Teoria da Resposta ao Item (TRI), com detalhamento dos modelos logísticos para 1, 2 e 3 parâmetros, e aplicação dos dados coletados no software BILOG.

A metodologia Teoria da Resposta ao Item proporcionou as análises das Curvas Características dos Itens e as descrições numéricas dos dados.

A dificuldade dos itens que constituem os testes pode ser verificada através da tabela 22:

<i>Grau de dificuldade/ano</i>	2000	2001	2002	2003
Muito Difíceis	16%	22,5%	6,67%	7,5%
Difíceis	44%	20%	40%	22,5%
Pouco Difícil	28%	25%	13,33%	27,5%
Fáceis	8%	27,5%	26,67%	27,5%
Muito Fáceis	4%	5%	13,33%	15%

Tabela 22 – Percentual de questões em cada nível de dificuldade – 2000 a 2003

A Proficiência/Habilidade média dos respondentes, o total de examinados e itens, a média de acertos e o conceito do MEC/INEP é apresentada na tabela 23.

Tabela 23 – Habilidade/Proficiência média, total de examinandos e itens, média de acertos e conceito MEC/INEP

Dados/ano	2000	2001	2002	2003
Habilidade/Proficiência	-0,299	-0,312	-0,276	-0,295
Total de Examinandos	46	59	41	41
Total de Itens	25	40	30	40
Média de Acertos	10,52	16,63	13,10	14,80
Conceito MEC/INEP	C	C	C	B

Fonte: O autor

Analizando a Tabela 22, verifica-se que a dificuldade geral dos itens diminuiu no decorrer dos anos, este fato pode ser constatado já que a proporção de questões fáceis e muito fáceis aumentou, enquanto as proporções difíceis e muito difíceis diminuíram. Com relação à habilidade/proficiência média, a Tabela 23 mostra que não existiu grande variação nos anos estudados. Portanto, essa queda na dificuldade da elaboração dos itens do ENC/Provão pode, juntamente com outros fatores inerentes à avaliação externa, explicar o fato do curso de matemática ter obtido o conceito B no último ano avaliado, mesmo com as proficiências/habilidades dos respondentes no decorrer dos anos permanecerem mais ou menos constantes.

A cada ENC/Provão são feitas análises pelos professores do departamento e os mesmos acompanham os resultados obtidos com a finalidade de melhorar a qualidade do curso, a fim de atingir o conceito “A”, que é a excelência estabelecida pelos critérios do MEC/INEP.

Com todas estas informações e subsídios fornecidos pela TRI, os professores do Departamento poderão, com muito mais segurança, fazer um diagnóstico completo de cada item, podendo ser pelo assunto abordado, pelo grau de dificuldade encontrado pelos

respondentes, pelo grau de discriminação dos itens ou pelos erros mais freqüentes, etc. Pode-se melhorar a discussão e analisar o ENC/Provão como um todo, este é o grande ganho de qualidade que a técnica TRI traz para a avaliação.

A aplicação da TRI nos dados coletados do MEC/INEP foi a contribuição real deste trabalho para o Departamento de Matemática da UNICENTRO. A partir deste poderá ser feita a mesma análise para os demais cursos da Universidade, bem como de se fazer com que os procedimentos de avaliação cientificamente comprovados passem a ter um uso rotineiro dentro da IES.

Como sugestão para trabalhos futuros, tem-se a utilização da TRI no Exame Nacional de Desempenho Estudantil – ENADE, cujos exames serão aplicados no primeiro e no último ano de cada curso, onde várias outras questões na área da educação poderão ser respondidas através da TRI.

REFERÊNCIAS BIBLIOGRÁFICAS

ANDERSEM, E. B. **Conditional Inference in Multiple Choice Questionnaires**. British Journal of Mathematical and Statistical Psychology, 1973.

ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. C. **Teoria da Resposta ao Item: Conceitos e Aplicações**, 14^o SINAPE, Caxambu – 24 a 28/07/2000. ABE-Associação Brasileira de Estatística.

ANDRIOLA, W. B. Artigo: Utilização da Teoria da Resposta ao Item (TRI) para a organização de um banco de itens destinados a avaliação do raciocínio verbal. **Psicologia: Reflexão e Crítica**, Universidade Federal do Ceará, 1998.

ANDRIOLA, W.B. Artigo: Descrição dos Principais Métodos para Detectar o Funcionamento Diferencial dos Itens (DIF). **Psicologia: Reflexão e Crítica**, p. 643-652, Universidade Federal do Ceará, 2001.

BAKER, F. B. **The Basics of Item Response Theory**, University of Wisconsin, 2001.

BIRNBAUM, A. **Some latent trait models and their use in inferring an examinee's ability**. In Statistical Theories of Mental Test Scores, F. M. Lord and M. R. Novick (Eds.). Reading, M. A. Addison-Wesley, 1968.

BOCK, R. D. and AITKIN, M.; **Marginal Maximum Likelihood Estimation of Item Parameters: APPLICATION OF AN EM ALGORITHM**. Psychometrika, vol. 46, nr. 4, p. 443-459, University of Chicago, dez. 1981.

BOCK, R. D. and LIEBERMAN, M. **Fitting a Response Model for n Dichotomously Scored Items**. Psychometrika, vol. 35, nr. 2, p. 179-197, junho 1970.

BOCK, R. D. and ZIMOWSKI, M. F. Multiple Group IRT. In **Hand-book of Modern Item Response Theory**. W. J. van der Linder e R. K Hambleton Eds. New York: Springer-Verlag, 1997.

BOOMSMA, A.; DUIJN, M. A. J. VAN. SNIJDERS, T. A. B. **Essays on Item response Theory**, Springer, New York, 2000.

CARVALHO, E. S. A. **Concepção de Educação Profissional no Exame Nacional de Cursos.**– RAIES -54. 2001.

DEMPSTER, A. P.; LAIRD, N. M. and RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm (with discussion). **Journal of the Royal Statistical Society, Series B**, 1977.

EMBRETSON, S. E.; REISE, S. P. **Item Response Theory for Psychologists**. New Jersey: LEA, 2000.

FLETCHER, F. R. A Teoria de Respostas ao Item: **Medidas Invariantes do desempenho escolar**. Ensaio. 1994.

GULLIKSEN, H. **Theory of Mental Test**. New York: John Wiley and Sons, 1950.

HAMBLETON, R. K.; SWAMINATHAN, H.; ROGERS, H J. **Fundamentals of Item Response Theory**, Sage Publications, London, 1991.

INEP. **Acabou o Recreio. Brasília**, (n. 4), 20-Ensino Superior. São Paulo: Segmento, (n. 17), 10/12/2000. <http://www.inep.cob.br/saeb>.

INEP. Relatório – Síntese, **PROVÃO, Exame Nacional de Cursos**, Ministério da Educação, Brasília-DF, 2003.

LORD, F. M. **A theory of test scores**. Psychometric Monograph, New Jersey, 7. 1952

LORD, F. M.; NOVICK, M. R, **Statistical Theories of Mental Test Scores**, Addison-Wesley Publishing Company, New Jersey, 1967.

MEC. **Lei de Diretrizes e Bases**. Lei n. 9.394, de 20/12/96.

MISLEVY, R. J. Recent developments in the factor analysis of categorical variables. **Journal of Educational Statistics**, 1986.

MUÑIZ, J. **Introducción a la Teoría de Respuesta a los Items**. Madrid: Pirámide, 1997.

MUÑIZ, J. **Teoría Clásica de los Tests**. Madrid: Ediciones Pirámide, 2000.

NEYMAN, J. and SCOTT, E. L. **Consistent estimates based on partially consistent observations**. *Econometrika*, 1948

PASQUALI, L. Teoria da Resposta do Item – ITR: **Uma Introdução, Teoria e Métodos de Medida em Ciência do Comportamento**, p. 173-195. Brasília, DF: INEP, 1996.

PERRENOUD, P. **Avaliação: da excelência à regulação das aprendizagens** – Entre Duas Lógicas. Tradução Patrícia Chittoni Ramos. Porto Alegre: Artes Médicas Sul. 1999.

RASCH, G. **Probabilistic Models for Some Intelligence and Attainment Tests**. Copenhagen: Danish Institute for Educational Research, 1960.

SAEB, -**Sistema Nacional de Avaliação da Educação Básica**, <http://www.inep.cob.br/saeb>.

SOARES, T. M; PEREIRA D. R. M. Artigo: **Estudo de Critérios de Ajuste para Modelos da Teoria da Resposta ao Item (TRI)**, SBPO - Simpósio Brasileiro de Pesquisa Operacional, Campos do Jordão – SP, nov. 2001

TAVARES, Heliton R; Andrade D. F. Artigo: **Item Response Theory for Longitudinal Data: Item Parameter Estimation**, Departamento de Estatística da UFPA - 2003

TURIM, M, E, **Análise de Itens pela Teoria Clássica de Avaliação e TRI em Dados Reais do Ensino Fundamental**, Dissertação de Mestrado: FAFICOP/UNESPAR, 2003

VALLE, R C. **Teoria da Resposta ao Item**, Dissertação de Mestrado: IME/USP, 1999.

VENDRAMINI, C. M. M. **Aplicação da Teoria de Resposta ao Item na Avaliação Educacional**. Em R. Primi (org). **Temas em Avaliação Psicológica**. Campinas: IDB Digital/Instituto Brasileiro de Avaliação Psicológica (IBAP). 116-130.2002.

VIANNA, H. M. – **Testes em Educação**. Ed. Ibrasa Inst. Brasileira de Difusão Cultural Ltda. São Paulo – S.P, 1973.

VIANNA, H. M. – **Introdução à Avaliação Educacional**. Ed. Ibrasa Inst. Brasileira de Difusão Cultural Ltda. São Paulo – S.P, 1988.

YAMASCHITA, S. S. A Avaliação do Ensino Superior e o Exame Nacional de Cursos: Percepções do Corpo Docente de Instituições Particulares de Ensino de Administração de Empresas. **Caderno de Pesquisas em Administração**. São Paulo. v. 7, nr. 4. 2000.

WRIGHT, B. D., STONE, M. H., **Best test Design - Rasch Measurement**, Mesa Press, Chicago-IL, 1979.

WRIGHT, B. D. **Sample-free test calibration and person measurement**. Proceedings of the 1967 Invitational Conference on Testing Problems. Princeton, N. J.:Educational Testing Service, 1968.

ZIVIANI, Cilio, Artigo: Provão 2001 Psicologia, **Estudo sobre as Questões de Múltipla Escolha**, Universidade Gama Filho, Rio de Janeiro – RJ, 2002.

ANEXOS

EXAME NACIONAL DE CURSOS - ENC/PROVÃO 2000

Escolha Questão Objetiva (*1)	Gabarito (*2)
EABABAACBABAADBBAAABBACBB	1010000100000011000101000
BADBA AE CAAABECE DDCEB B	0000100001000000000010100
EBADCABCCCBAEACDCBAACCAB	1001000111000000010000010
EEBDAACEEBBCBBBCDABCBEBCC	1011100000000000000000100
BACDDAEEABCBEDCDCADEAABAD	0001000000101100000001110
BCBDAEACEACBEABEDBBBEBAD	0011100100101000000100110
EBCEAEDEEBCBAACBDAABABACE	1000101000100001000100000
EAADCBAAEADACECDABCACCBCE	1001000000000000000000000
EAADADBEBCBAABEDCDBECDDD	1001110001000000010100000
EACAAABBEBCBEDDCBABDADCBC	1000100000101110000000000
EBEAEEEACEBEADDDCDDCEEB	1000100001001010010000000
BAEEAACAABBBBDDCDABBBCAAD	0000100000000010000100010
EABDAAACCCBCEAAEDBABADDDA	1011100111001000000100001
BDBBAAACEBEDADDDDABAEBCEB	0110100100000110000000000
ECDACAEECEDBBAECCAEACBAEE	1000000010000000001010000
EECDAAACEABADDBBCBDBBEEAB	1001100100000101000100010
BACCCACEEADCBACEDEBBAAAAC	00000000000000000000101010
EACECDCBCBCCAABCBBCBCCAE	1000010010000000000110010
ACEEBCCBABB CABDCCAACDBABC	0000000000000010000000000
BDBECBADDEEACEECBABBCDBDB	0110000000000000000110100
EDDDEDADCEBDBBEEDBACAECAB	1101010010000000000000010
BACBACBEEBBEEAABDAEABECED	0000100000011001001000000
EEAEABACEBABAACEDAABBACBB	1000100100000000000101000
EBCCAACECECB EADCCCECBDEA	1000100010101010000000001
BDBAACCECBAAACCCAEBACBCD	0110100010000000001100100
EDCBAEADCBAAACABBAACBABECA	1100100010000001100100001
EDEDCABEADEABACCEADECBCEB	1101000000000000000010000
EABBBAADDDCBADCB BBBBABEB	1010000000100010000101100
BAACADCBDBCDBAEDBDBCAEABE	0000110000100000000000000
CCCBDAEAE EBCBECCBCCACBDB	0000000000000000000000100
EBAEECDEABCACBBBBAEABABAE	1000001000100001001001110
EDCCADBCCCCAADABBBEBCADCB	1100110111100101001111000
EEBDDEAEDCBAADCEDAECBACCE	1011000001000100001001000
BDDBCCAAEBCADACCDCEBBCCCE	0100000000100000011100000
BACADAACEEBDAADDAECECCDBAE	0000000100000010111010110
EDBECDBACBEBDBBCBABAABDAE	1110010010000000000000010
EABBAECBEBBDECDCEABBBBDBD	1010100000001010000100000
BECABAAECCBAAACCDABBACEAE	0000000011000000000100010
EEEEBDACEEADDADEEAECBAACD	1000010100000010001001000
CACCEEAEABBBBABEDABBADABA	0000000000000000000100001
EABDCACEDCCABCDEBCACEECAB	1011000001100010010000010
BCCDDCAECEBDBADCBBCACDCD	0001000010000010010000000
DDBACBBEBEBABAABDAECDDBB	0110000000000001001000000
BACBAAEBBBAEBACEAADDABABE	0000100000010000100000000
ECDDBDCADBCADBCCAABABCEAA	1001010000100000100000011
DD BCBABEEADAABBABCACEECA	0100000000000001100010001

(*1) - Respostas - Gabaritos Respondentes

(*2) - Respostas Objetivas - Acerto (1) Erro (0)

EXAME NACIONAL DE CURSOS - ENC/PROVÃO 2001

Escolha Questão Objetiva (*1)	Gabarito (*2)
AECADEAECEBCCCEAECEBDBEBECABDDCBEDABBCA	1010101100001001000000100000010000100000
CBAEDAECEBCCBEACAAEDCCAADEACDAECDDAEACCD	01000011110011010010010000000010100101001
ECAEBAECEBCCABEEBAECDCBABCACCBECDBCCB	0000000110000100100001000000000000000000
EBDBDBEEBCCDBACBBABABCEDBBADBADCBDBBEED	0100100100100100111100000100011000010011
ACCEDBAEBCCECAECABABDCDCACBACBAEBCBAACE	1010101100101000011101110000001010011100
EACCEAECEDDCBEDAAAADCBBAEABAADCAEBD	0011001110001100001001100000000000101001
EBAEAEAEBAEBDBACCECEDBCABDCADDECEDDBBCCA	0100001100000100000000000000110000110000
ABBECAEBBCBCECBADDDCAADCAEAAECADBBCDC	1100011100101100111001100000000000110000
AACDDAECEBDCBAAABAAACDDEABAAABCBDCBEDD	1011101110001101011001101000000111100001
EAACBBAECCDCDDEBBABDCDBEBAADACEECBCECB	0001001110101010111101100000010000010000
BABECCAECBCCBCEEDAAACDECDCCDECCDBBBAC	000001111000000001001100000100100110000
ABCCCCAECDACCBDABABACDCABEBAABCEBCBAEED	1111011111001100011101110100001001011011
BA*CBDECCABABACABACACBBACACACCBDEEEE	000110011000010001100100000000001010010
DAEEDEEE*ECADEBDCADDDCD*DBCCADCDCEAD	0000100100110000001000000000000000100001
AABBAEAEBCADCCCAECABDECAAAEAABBCBDAEAE	1000001100001001001100000000001000100110
BEEAECEBBDABBEAABEDADACBEBBCDEBCABBBBAC	0000010101000001100000001100101011010000
EBCECACADCECBEAABABDDACEDECBABDCDBADEE	011101100111101011100011000001000111010
AAECBAEBDBCCBECAEEADCCBCCACABDBDBACCD	1010001101001100000001000000001000111001
EBADECBEBDADCBECABADDDAABDBCDABCCBCADDDE	0100010101001100011000000000101001000000
ABAEBADAEBEECEDECBEBDBCDDBABAEACBBBEBE	1100000000011000010100000000000010010000
ABCABCAEEAAABCEBCDADDACDDCBBAABCCAEABA	111001110000000001000000000000000000100
AADEBEECBBAEAECCABDACEBCCCAEADDDABEAB	10001001100010000011000001000000000010000
ABCEBAECCBDBEEEBAAADCDDBCCAEECECBCEEE	1110001110000100011000000100000100010010
AAECBEDCAEDCBECBBACACBEBDC*DABBCABDECEA	1010000010001100111001000000101000000010
ABCCABEEBDBECCCBCEADADBCABEACAACBBBEAECE	1111000101011000101000010100000001001000
ACCEDEEECDCECBACADADCDCBECAADAEEEDBADDD	1010100111111100001000001010010100011001
AACDDCAECBBDBECABBBBCCBCECACBEBDAADCE	10111111100001000101010100001000001101000
AEBDBAECCACDCDEAAAAAEDBBEAAACDACBBAECA	1000000010101001001000100100000010011000
ABADECBBDACDACACABACDABAEBAACCBABBCAA	11000000000010000011011000000000001010000
BAEACEDECCBDBCAEDBABBABADDCBABCBBADBDABE	000000001100000000111000000000001001000000
AACEDBACCBBCCBADBBEBAEAEAEBCABDBDBABDE	1010101010001100110101000000000001111000
ABCDDEDACBBBDBEBCAEAEBCDACAACCEBBBBD	11101000100001101010000000000000000010000
EAEDDDEBCEADEADABCCADCCDADDACDCADAEEDE	0000100010000000100001000000100000101000
AABEECEBEBCEDEBDCADDBBCCCEAADECCECBABDA	1000000000100000001000000000010000011000
EAEAECEEBEBAEAEAAADAAAEAAEACAABBAABAA	000000001100000000010000000000000010010100
EAAEAEAECCBCEEBABBDDECEDCBCDECDADBBECD	0000101110001100011100011000010000110001
AAAEAEACCEBDBEEAEEDDBCCAAEABEAEDEAAACCA	10001010100001000000000100000000100001000
ECBEADAACCBCEBABBDADEDECBACBDACAABCCB	0000001010000001110000101000000000010000
BADDDADECDCCBDCDECCBDECBCCDDABCDADDEC	0000100111101110000000100100010000101010
EEBAEAEDEDECEBAEDEAABABDBACAEECCBBCAEBE	00000000010110000100000000000000001001000
AAAEDEACEBCACDEBBABDADDDBACECCCBDDCBA	1000100010000010111100100000000000000000
EBAEDEACEDCBACEDEAECDCEBEECAADECCDBBBA	0100111111100000000011000000001000000000
AAEEADEEEBCEDBEDCAAEAECEDEEDABDACEDEABAACA	1000000100110100001000001000010000011100
EACBDDABDCBACBBABAEADDEBEEAADCAAEABABAE	0010101011000100111000100000010010011000
ECEEDBAEEDCCBEAABBDCEBDBCCDACEEDDDDBDA	000010110110110101010011000000100100100000
EAAEEDABBAECDBABEBAABDCAEBCEDACBBDBBECD	0000001000000100011000000000010001110001
AABADCEDEBABCBECDACBBDBDEBABDBCDADABCD	1000110100001100001000100000000000001001
AAAEBEEDCCAACDAECDABDACEBEEAEACADDBBCDD	10000000100010000011000000000000001010001
DBEAAEAEBDECCCEABEDDDAECEACAACDBABEBBD	0100001101001000010000001000000001010001
EDEEEAEAEADBCABEBBDECBCEDCBABBACDBBCCC	0000001100000000010011011000001010110000
EDADCBD* BDBDAEACABDDABBCABADDADBBBBD	0000000100000000001100100100000010010000
AAECBDEBABBCEACBAAACBAAAECDAAADBAEEDDB	1010000100001101011001000000100001000000
EADCDDEDCAEBCDDEDAEDDDDBCEEABACDCEEBCCCC	00011000100010100000001000000000000010000
CCACAADCBAADACAEBCEDECEBBBAABCBADADCCD	0001000000000000000100011000110000000100001
EDEBEDEEBEAAEDBCEBDBCCBBDCCDACCEDDBDCA	00000001000000000100101000000100000100000
ACCCDEBECBACBAEADABABDDCEDCDABABDDDBDEBE	1011100110100000001100100000101000110000
D CAACDAEEABBCEDBABAECBAABBEDDBDE	0000000010000000010010000000000000100000
ABDADBADCEDDCBACBCEABCEBDCBEABDBBDBACDE	1100101010001100100001000000001001111000
AADEEEBCEEBEB*AEBCCEDCBABDACACEEEBBDEDA	10000000000100001000010001000000100010000

(*1) - Respostas - Gabaritos Respondentes

(*2) - Respostas Objetivas - Acerto (1) Erro (0)

EXAME NACIONAL DE CURSOS - ENC/PROVÃO 2002

Escolha Questão Objetiva (*1)	Gabarito (*2)
CBCDECECAEABEBAAAEDDDAECACBBBA	00010000000010101000011001101
EDCBBBACCEBECCECEBDDBEADEBBBA	000001000010001000000110111101
CAEBCAAECCEBACDECCCCACDECCEBCCA	001010000011001000000011011001
EAEBCCCCDCACCEBACDDCDBEDCEDBBA	001010001100000100010010010101
EBAABABEADADAECCEBCADDDCCAABA	000000101000000001011000000001
EBAEBABCBDDDEABACDEDCBDDACDEAA	000000100000000100000000000001
DCABAAADCABACDDCADEDCEBEAAAA	11000000001100000000000010001
CCADCBDAAABCDEEACEEDDEAECEEBCEA	010111000010101101000100011001
CCABCBECCABADAAACDBDADBBACACBA	010011000011100100001000000001
DDABCBCEDBBADBEABBEBCCEACBACCB	100011001011111100000101000000
EBCBAAEDEEBBADEAAADDDACEAAADCB	000000000010001100000000000000
DCEDCBACEBABBAAABDAECEEEBEDDDA	111111110011010100100110010011
DCDDEABACABADDEADBDECDADCAAABA	110100110011101110000000000001
BCECEAECEDEBAADDEEADCEEBBAAAE	011000001011001000000110000000
DDDBDDDDBBEBCDEBAEABCDEBDECACA	100000000000001001100010110001
DCADCBBADCEBCABEABAADAEEDEBBAA	110111111010011100101110111101
DDABABAABABBACCAEDBCCEABCCBCA	100001010010000001000010000101
DEABBBACEEAAADAACDBDCDAEEDACAA	100001000001000100000000000001
DCDDCBBCBCBCDDECDCCBCECEEBCCA	110111100110101010000101011001
DCADAAEABEBCAEEABACBDDACBEBABA	110100010010001100000001011001
CAEDCBEBABBADACCAEDBAACEDBDCB	001111100010000000000001001000
ECABCBEAAAEBAEAAABADDAADAECCEA	010011010000000100100000010001
DCAACBBEDABCEDEEBBDEBCCACBBAA	110011101010001000000001001101
AEB*BEBEABEAE*A**DEED*EB****	000000100001001000000100000000
CABECEBADCBBDDEEACDEDCDEDEDAABA	000010111110101100000010110001
ACBDCEBCEDBBDDDEEEDDEEABCEDEBCBA	010110100010101011001000111001
DBADCBCCAEBEADDEEAABDCDEEAAADCA	1001110000100010000000010000001
CBADCABEBEEDAEABEDADCCABEEAAAA	000110100000000000100000010001
DCADCCBCBABEDBEADEBBDDDEEDEBCDA	110110100010111111000010111011
CECBACCEAEDAEDCCBDEBCDBBBEABBA	000000000001000000000000010101
DCADCAEEBEBBEDDABEADDCBEEBDAA	110110000010000101100000011001
DCEECBBCEABDADEAADDDDECBAEBBA	111011100010001100000011000101
DBDDCBDCDABDADEADBADDEEDAABAAA	100111001010001110100110001001
DBCABADEDEEBADBCCBADCEADBAABEA	100001001000000000100100000101
DDADCBBCCCEBDBDEABCEDCDAEBABCEA	100111100010001100000000001001
DBEDCBCEEEBBABEEACEECDABAAADBA	101111000010011000000000000001
CAACCBACDEBEADDCCDABDDCEAADDDBA	000011001010000000100000000001
DBEBEBCECAEAAAEBDDDDADDCEBEEA	101001000001001000001001011001
DDDEBCECEEEAEDEAADABEDCEEEAEBA	101100000001001100100000010001
EEACBDEAEBADEBABEDDDADEAABEA	001011000000001000000000000101
EBEBEDDCCEABCDEAAAEDCEDCBCDAA	001000000000001100000010000001

(*1) - Respostas - Gabaritos Respondentes

(*2) - Respostas Objetivas - Acerto (1) Erro (0)

EXAME NACIONAL DE CURSOS - ENC/PROVÃO 2003

Escolha Questão Objetiva (*1)	Gabarito (*2)
BAEEEABDDDCBADACBBBCDABEAEADEBBBAEACEAA	0110000000100001010000001100000000000000
EAEDBEDDACCEACDEABADBCDAEBECCBCEADBECED	1110101001100010100001001110010100110111
AAECDBBDDCBEEBAEBDADABBACBDACBAEBEBEBE	0111100100100000111000001010110000011000
ABBBAAAAEABAAECAEEAAEAAEABECECBCEABBD	0000000000000001000000100100010100001001
DAEBEBAEDDAEEACBAEBABBCBEDDCBCCDCEBCCCB	0110000000000010000001000100000110010100
EADBABEEEAEBEACCCDDDBCEBDAAADDBDEDCCCD	110000100100000000000000010100000000101
AAEDDEECDBCBCAEEDBDBCEEEAEACABAEDADDCBAA	0110101000100000011010001101101010100000
EAEDDCEBDACCBBCAEBDADCECADEAEACABDBDDCCEE	1110101101100110111000011000011010100110
EAEDCBEBDDBBBEBEEDCBDEAEACADBCDCCDD	1110001100000100100111001101111000100101
EAECDEEBDACABBBABEDADBEDABBAACECDBDCCBC	1111101101110100111001001010110110100100
CAEAABDBAECEAECEDECBCEEEAAADDADBDCCBEA	0110000100100001101011000100100010100010
BAEACCEBABCABEBEBEADBBEDCAAECCEADACEBB	0110001100110100110001000000010000100000
AAEAABACDACAEEAAEAECEEEAEDECAEBDEACEAD	0110000001110000100000001110011000000001
ABEDDBCDAEBCECBEEADBA AEAEACCDADDEEBAD	0010100101000010100001101100110000001001
EAEBEAEDDBCBAEECEBDAEBDAEAABAABDAAECBBE	1110001000100001111000001100001010000000
AAEDAEDAABCBEBCAACAAEACDEAEDBAEDDBBCDBB	0110000000100101000000110100001010010000
AAACACDDABCDABEBACBCEBEDEBADEEACBDEACDEB	01010000001001000000001000000001100000010
EAEAEBEDDDCBEBEBCADACBAEDECEDAECCBDCDEEE	1110001000100100001011100100000100100010
DAEBABABAABDEBCDABEACBADAEADADEAECECBAB	0110000101000110010011101100000000000000
EBEDDBEDEBCBBECEBBADBCDBEAACABDCDBBBED	1010101000100101110001000100111010111011
AAECACABDACDABAECCBBBDADCBCBAACCEECBCD	0111000101100100100001001000001000000001
EAECDDBEABBEBCCEBDDDBACAEBBCACDEDEBCDD	111111110100011111110111111011111101101
DAECDEEBDDCCDBEBEBECABACDEAEBDDEEADBBAD	0111101100101100110001110100000000111011
DADCAEBEDACCABADABECCBBDEEEDBCDABADCEE	0101000001100100010011000100000000000010
EAECBBECCBDBBCEBCADBDDDECDBAECDADCECED	1111001000000111110001000100000110100001
EAEDDBEABACBAAEEEBDADBDDBBDACDCDAECBBAE	1110101100100100111001000010110110001000
DAECDBEABDAABBDDEEBDBDEAEAAECCEAADECAED	0111101010010100110001001100010000100011
EAEDDBEBAACBBBCBEBDADBDAAEEAACCCDECBACE	1110101101100110111001001100110111010000
ACEDDDAEDCCCCBCDABACDEBDEEAAEEBCACBCCDAC	0010110000100110010000000100000100000000
EAEECECAEDBEECBDECCDBEAEDDDDAEBADDBAEC	1110001000000010000001001100001000101010
DBECDAAEAACBEACBDCBAACDCABACAABDEEBBCA	0011101001100011011000110000111000001000
AAEDDDEEDABADBBEBBDCBDEEDCACACDABECDBD	0110111001011100011011100101111110000001
AAEBDDEBABABAECECEADBECEACEAACDCCECDAA	0110111100000001100001000101001110000000
ABECCABBADAEBCEDEACCACEACBCAEDDEBDCCDEC	0011000100000110000010001011100000100010
DAECDACEDCEBECEEAECCEBDEEBDEBDDEADBBBCD	0111100000000000000011000110000000111001
AAEDBAEDDACCEAEEDBCDBACDEADEBEDDDDACDDA	0110001001100000111001110100000010100000
AAECACBDDEABAEAAACDCEBDEBACACBEBBABC	011100000000000000001000000100111000011000
BAEEDDEBDBEABEBCBDEBBACAEBAEEACDABBDABE	0110111100000001111001111110100000011000
AAEADBCCDEBCEBCEADABACAEAEBCDEAEDCEBB	0110100000000101000001111100000000000000
AAEDEBEECBEDABBEBCBDDAEACEEEDACACCCBC	0110001000000100110001001101000000000100
AACBACBDDAABABEEEBCCDCAAEBDAAEDCBAACACA	0100000001000100110000001110100000000000

(*1) - Respostas - Gabaritos Respondentes

(*2) - Respostas Objetivas - Acerto (1) Erro (0)